DUKE MATHEMATICAL JOURNAL

EDITED BY

LEONARD CARLITZ   DAVID VERNON WIDDER

JOSEPH MILLER THOMAS
Managing Editor

Volume 9, Number 4
DECEMBER, 1942

DUKE UNIVERSITY PRESS
DURHAM, N. C.

DUKE MATHEMATICAL JOURNAL

This periodical is published quarterly under the auspices of Duke University by Duke University Press at Durham, North Carolina. It is printed at 8 North Sixth St., Richmond, Virginia by the William Byrd Press.

Entered as second class matter at the Post Office, Durham, North Carolina, with an additional entry at the Post Office, Richmond, Virginia.

The subscription price for the current year is four dollars, postpaid; back volumes, five dollars each, carriage extra. Subscriptions, orders for back numbers, and notice of change of address should be sent to Duke University Press, Durham, North Carolina.

Manuscripts and editorial correspondence should be addressed to Duke Mathematical Journal, 4785 Duke Station, Durham, North Carolina. An Author's Manual containing detailed information about the preparation of papers for publication will be sent on request.

Authors are entitled to one hundred free reprints. Additional copies will be supplied at cost. All reprints will be furnished with covers unless the contrary is specifically requested.

The American Mathematical Society is officially represented on the Editorial Board by Professors Murray and Ward.

QUASI-ANALYTICITY AND PROPERTIES OF FLATNESS OF ENTIRE FUNCTIONS

By S. Mandelbrojt

Introduction. We shall denote by $C\{M_n\}$ the class of all infinitely differentiable functions which are defined in a closed interval $[a, b]$ and which have the following property: To each function $f(x)$ of the class there corresponds a constant $k = k(f)$ such that $|f^{(n)}(x)| < k^n M_n$ $(n \geq 1, x \in [a, b])$.

The class $C\{M_n\}$ is said to be quasi-analytic if from the two properties $f(x) \in C\{M_n\}, f^{(n)}(x_0) = 0$ $(n \geq 0, x_0 \in [a, b])$ it follows that $f(x)$ is identically zero.

Theorem A. *A necessary and sufficient condition for $C\{M_n\}$ to be quasi-analytic is that*

$$\int_1^\infty \frac{\log T(r)}{r^2}\,dr = \infty,$$

where

$$T(r) = \underset{1 \leq n}{\text{l.u.b.}}\ \frac{r^n}{M_n}.$$

This is Denjoy-Carleman's theorem [1]. The particular form given here is due to Ostrowski [3].

If there exists in $C\{M_n\}$ a function $f(x)$, not identically zero in $[a, b]$ and such that $f^{(n)}(x_0) = 0$ $(n \geq 0)$, there exists also a function $\varphi(x) \in C\{M_n\}$ not identically zero and such that $\varphi^{(n)}(a) = \varphi^{(n)}(b) = 0$ $(n \geq 0)$ [6]. Thus Theorem A may be stated in the following manner:

Theorem B. *A necessary and sufficient condition that there exist a function $\varphi(x) \in C\{M_n\}$, not identically zero and such that $\varphi^{(n)}(a) = \varphi^{(n)}(b) = 0$ $(n \geq 0)$, is that*

$$\int_1^\infty \frac{\log T(r)}{r^2}\,dr < \infty.$$

The present author has introduced the notion of quasi-analyticity for functions

Received February 12, 1942; presented to the American Mathematical Society October 31, 1942.

integrable L (we shall say integrable for integrable L). Let $f(x)$ be integrable in $[-\pi, \pi]$. For $x_0 \varepsilon [-\pi, \pi)$ let us put

$$\varphi_r(x_0, \alpha) = \int_{x_0}^{x_0+\alpha} | f(x) | \, dx \qquad (\alpha > 0, x_0 + \alpha \varepsilon [-\pi, \pi]),$$

$$\limsup_{\alpha \to +0} \frac{\log(-\log \varphi_r(x_0, \alpha))}{-\log \alpha} = \rho.$$

The point x_0 is said to be a right-hand mean-value zero of exponential order ρ for $f(x)$.

If $x_0 \varepsilon (-\pi, \pi]$, and

$$\varphi_l(x_0, \alpha) = \int_{x_0+\alpha}^{x_0} | f(x) | \, dx \qquad (\alpha < 0, x_0 + \alpha \varepsilon [-\pi, \pi]),$$

$$\limsup_{\alpha \to -0} \frac{\log(-\log \varphi_l(x_0, \alpha))}{-\log(-\alpha)} = \rho,$$

the point x_0 is a left-hand mean-value zero of exponential order ρ for $f(x)$.

The following theorem was proved by the author [3]:

Theorem C. *Let $f(x)$ be integrable in $[-\pi, \pi]$, and suppose that $x_0 \varepsilon [-\pi, \pi]$ is a right-hand (or a left-hand) mean-value zero of exponential order ρ for $f(x)$. Suppose that the Fourier series of $f(x)$ is of the form*

$$f(x) \sim \sum (a_{m_i} \cos m_i x + b_{m_i} \sin m_i x), \tag{1}$$

where the index of convergence of $\{m_i\}$ is $\sigma < 1$. If

$$\rho > \frac{\sigma}{1 - \sigma},$$

$f(x)$ is zero almost everywhere.

A new proof of this theorem based on the theory of Fourier transforms, but closely related to that given by the author, was published later in a joint paper by N. Wiener and the author [4]. They have also proved that, from a certain point of view, this theorem cannot be improved.

They proved in fact that, given any $\sigma < 1$, it is possible to construct a continuous function $f(x)$ of the form

$$f(x) = \sum a_{m_i} \cos m_i x,$$

not identically zero, for which $x = 0$ is a right-hand mean-value zero of exponential order

$$\rho = \frac{\sigma}{1 - \sigma},$$

where the index of convergence of $\{m_i\}$ is σ.

Recently B. Levine and M. Lifschetz [2] published a new proof of Theorem C. They gave also a profound generalization of this theorem. If the Fourier series of $f(x)$ is of form (1), let us denote by $N(t)$ $(t > 0)$ the number of positive integers m_i $(m_i \neq 0)$ such that $m_i < t$. Let $A(\alpha)$ be a function defined and positive for α positive, increasing to ∞ when α tends to zero, and such that, for α sufficiently small, $A(\alpha) \leq -\log \varphi_r(-\pi, \alpha)/\alpha$. For α sufficiently small $\alpha A(\alpha)$ can then be expressed as a function of $A(\alpha)$; suppose that $\alpha A(\alpha) = \omega[A(\alpha)]$. Further, set

$$\zeta(r) = 2 \log r + 2r^2 \int_0^\infty \frac{N(t)dt}{t^2(t^2 + r^2)}.$$

The theorem of Levine and Lifschetz can be stated as follows:

THEOREM D. *If*

$$\liminf_{r=\infty} [3\zeta(2r) - \omega(r)] = -\infty,$$

$f(x)$ is zero almost everywhere.

Their method consists in reducing their problem to the problem of determination of an entire function by its values in a sequence of points.

If we assume that, for a sequence of values $\alpha \to +0$, and for ϵ any positive number,

$$\log \varphi_r(-\pi, \alpha) < -\alpha^{-\rho+\epsilon},$$
$$N(t) < C + C_1 t^{\sigma+\epsilon},$$

the first inequality holding for α sufficiently small and the second for t sufficiently large, Theorem D reduces to Theorem C.

Levine and Lifschetz proved also the following result:

THEOREM E. *Let $f(x)$ be integrable in $[-\pi, \pi]$; let, for $\alpha > 0$ sufficiently small,*

$$\log \varphi_r(-\pi, \alpha) < -\alpha^{-\rho+\epsilon};$$

and let the coefficients of

$$f(x) \sim \sum (a_n \cos nx + b_n \sin x)$$

be divided into two groups $\{a_{n_k}, b_{n_k}\}$ and $\{a_{m_i}, b_{m_i}\}$, where the exponent of convergence of $\{m_i\}$ is $\sigma < 1$.

If

(1) $$\rho > \frac{\sigma}{1-\sigma},$$

(2) $$|a_{n_k}| + |b_{n_k}| < e^{-\gamma(n_k)},$$

where

$$\frac{t\gamma'(t)}{\gamma(t)} \geq \sigma_1 > \sigma \qquad (t > t_0),$$

$$\int_1^\infty \frac{\gamma(t)}{t^2}\, dt = \infty,$$

then $f(x)$ is zero almost everywhere.

The authors suppose, therefore, that the function $f(x)$ and its coefficients with indices m_j satisfy the conditions of Theorem C and that the coefficients with indices n_k satisfy the conditions of a theorem of de la Vallée-Poussin [3] which is stated as

THEOREM F. *If $f(x)$ is infinitely differentiable in $[-\pi, \pi]$, if $f^{(n)}(-\pi) = f^{(n)}(\pi) = 0$ $(n \geq 0)$, and if all the Fourier coefficients satisfy* (2), *$f(x)$ is identically zero.*

As a matter of fact the present author gave previously [3] a theorem in which he did not suppose that $f(x)$ was differentiable, in which only a part of the coefficients satisfied a condition of type (2) (less restrictive) and in which the function $f(x)$ being *zero* in an interval implied that $f(x) = 0$ almost everywhere. The author also generalized Theorem F and proved its converse.

The purpose of this paper is to give a general theorem which may be regarded as a theorem concerning quasi-analyticity as well as the behavior of entire functions on a set of points. This theorem will contain as corollaries the parts of Theorems B, C, D, and E, modified in a certain way, which give sufficient conditions for quasi-analyticity. For instance, the functions will be supposed to have the same property at *both* end-points of the interval (at the right of $-\pi$ and the left of π) or, what amounts to the same thing, on both sides of a point $x_0 \in (-\pi, \pi)$. This theorem relates an integrable function to an entire function. Whenever an entire function is chosen there results a theorem on quasi-analyticity, and when a suitable integrable function is given, a theorem concerning entire functions follows.

The author believes that it is possible to give a still more general theorem which would also contain known gap theorems on the distribution of singularities of Taylor series.

The principal theorem. Let E be a point set in the open xy-plane. Denote by E_x the projection of E on Ox, and by I the smallest interval I_1 such that $E_x \subset I_1$. The interval I can be open, semi-closed, or closed. For every $x \in E_x$ we shall write $y_E(x) = \underset{(x,y)\in E}{\text{g.l.b.}}\, y$.

Consider the set of convex functions defined in I, where the function which is identically $-\infty$ belongs to the set.

There exists in I a convex function $C_E(x)$ with the properties:

(A) *For every $x \epsilon E_x$, $C_E(x) \leq y_E(x)$.*

(B) *Every function $\varphi(x)$ convex in I, such that for each $x \epsilon E_x$ $\varphi(x) \leq y_E(x)$, satisfies the inequality $\varphi(x) \leq C_E(x)$.*

The existence of $C_E(x)$ is well known, but it is of importance for this paper to notice that this function can be defined in the following manner.

For $t \epsilon T \equiv (-\infty, \infty)$ set

$$A(t) = \underset{x \epsilon E_x}{\text{l.u.b.}}\ (xt - y_E(x)).$$

Then

$$C_E(x) = \underset{t \epsilon T}{\text{l.u.b.}}\ (xt - A(t)).$$

If there exists a point of E below each straight line not parallel to Oy, then $A(t) \equiv \infty$, and $C_E(x) \equiv -\infty$. However, each function satisfying (B) is then identically equal to $-\infty$. Therefore, in this case, (A) and (B) are satisfied. In the contrary case, there exist two constants t_0 and k such that $y_E(x) \geq k + t_0x$, $C_E(x) \geq k + t_0x$ ($x \epsilon I$), and $C_E(x) \leq y_E(x)$ ($x \epsilon E_x$), where the inequality of convexity

$$(x_3 - x_1)\ \underset{t \epsilon T}{\text{l.u.b.}}\ (tx_2 - A(t)) \leq (x_3 - x_2)\ \underset{t \epsilon T}{\text{l.u.b.}}\ (tx_1 - A(t))$$
$$+ (x_2 - x_1)\ \underset{t \epsilon T}{\text{l.u.b.}}\ (tx_3 - A(t)) \qquad (x_1 < x_2 < x_3,\ x_1, x_2, x_3 \epsilon I)$$

is obvious. If now $\varphi(x)$ satisfies the hypotheses of (B), then obviously

$$A_1(t) = \underset{x \epsilon I}{\text{l.u.b.}}\ (xt - \varphi(x)) \geq A(t),$$

$$\varphi(x) \geq C(x) = \underset{t \epsilon T}{\text{l.u.b.}}\ (xt - A_1(t)) \leq C_E(x).$$

But if τ denotes the upper right-hand derivative of $\varphi(x)$ at a point $x = x_0$, which is not the right-hand end-point of I (when I is closed on right), we have, for $x_0, x_1 \epsilon I$: $(x_1 - x_0)\tau \leq \varphi(x_1) - \varphi(x_0)$. Therefore, $A_1(\tau) = \underset{x \epsilon I}{\text{l.u.b.}}\ (x\tau - \varphi(x)) = x_0\tau - \varphi(x_0)$, and $C(x_0) \geq \varphi(x_0)$. Thus $C(x) = \varphi(x)$ for every $x \epsilon I$ and (B) is proved.

We have proved, in passing, that a necessary and sufficient condition for a function $f(x)$ to be convex in I is that there exist a function $A(t)$ $(-\infty < t < \infty)$ with $f(x) = \underset{t \epsilon T}{\text{l.u.b.}}\ (xt - A(t))$. If $f(x)$ is convex in I, this equality holds with $A(t) = \underset{x \epsilon I}{\text{l.u.b.}}\ (xt - f(x))$.

Consider a sequence of pairs of numbers $\{(\alpha_n, \beta_n)\}$ such that α_n increases to ∞ and β_n is either finite or equal to $+\infty$. Denote by E the set of points (α_n, β_n) for which $\beta_n < \infty$. We shall call the *base* $\Pi(x)$ of the sequence $\{(\alpha_n, \beta_n)\}$ the

function defined as follows: If E is empty, $\Pi(x) = \infty$ for every x; otherwise $\Pi(x) = C_E(x)$ in the smallest interval containing E_x, and $\Pi(x) = \infty$ for the other values of x.

We shall say that $\int^{\infty} \Pi(x)e^{-x}dx < \infty$, if the integral $\int_a^{\infty} \Pi(x)e^{-x}dx$, for a sufficiently large, either converges or has the value $-\infty$.

We proceed with the proof of the principal theorem.

THEOREM. *Let $F(z)$ be an even entire function, not identically zero, and set $M(r) = \max | F(z) |$ $(| z | = r)$. Let $f(x)$ be an integrable function in $[-\pi, \pi]$, different from zero in a set of positive measure. Define*

$$I(\alpha) = -\log \int_{-\pi}^{-\pi+\alpha} (|f(x)| + |f(-x)|)dx \qquad (0 < \alpha < 2\pi),$$

and for $I(\alpha)/\alpha$ finite and sufficiently large, set $I(\alpha) = \omega(I(\alpha)/\alpha)$. Denote by a_n, b_n the Fourier coefficients of $f(x)$, and by $\Pi(x)$ the base of the sequence

$$\{(\log n, -\log | F(n)(a_n + ib_n)|)\} \qquad (n \geq 1).$$

If

$$\text{(I)} \qquad \lim_{r\to\infty} \frac{\log M(r)}{\omega(r)} = 0, \qquad \lim_{\alpha\to+0} \frac{I(\alpha)}{\log \alpha} = -\infty,$$

then

$$\text{(II)} \qquad \int^{\infty} \Pi(x)e^{-x}dx < \infty.$$

If

$$\liminf_{x\to\infty} \frac{\Pi(x)}{x} = k < \infty,$$

$\Pi(x)$ will be, by its definition, convex for x sufficiently large, and therefore for x sufficiently large $\Pi(x) < (k + \epsilon)x$ $(\epsilon > 0)$; (II) will then be satisfied.

Consider the case when

$$\text{(2)} \qquad \lim_{x\to\infty} \frac{\Pi(x)}{x} = \infty.$$

Let us denote by $C(x)$ $(x \geq 0)$ an arbitrary convex function satisfying the relationships

$$\text{(3)} \qquad C(x) \leq \frac{\Pi(x)}{2}, \qquad \lim_{x\to\infty} \frac{C(x)}{x} = \infty,$$

and set

$$\text{(4)} \qquad B(t) = \underset{0\leq x}{\text{l.u.b.}}\ (xt - C(x)).$$

As we have seen above

$$C(x) = \underset{t \,\varepsilon\, T}{\text{l.u.b.}}\, (xt - B(t)). \tag{5}$$

From the definition of $B(t)$ it follows that

$$\lim_{t=\infty} \frac{B(t)}{t} = \infty, \tag{6}$$

and (5) may then be written

$$C(x) = \max_{t \,\varepsilon\, T} (xt - B(t)) = xt_x - B(t_x) \qquad (x > 0), \tag{7}$$

with $\lim_{x=\infty} t_x = \infty$, the function $B(t)$ is a non-decreasing function of t.

If we set

$$c_n = e^{-B(2n)},$$

it follows from (5) that

$$c_n r^{2n} \leq e^{C(\log r)} \qquad (n \geq 1,\ r > 1).$$

If $2p$ is the smallest even positive integer greater than $t_x > 0$ $(x > 0)$, we see by (7) that for large values of $r = e^x$

$$e^{C(\log r)} \leq c_{p-1} r^{2p}.$$

It follows from (6) that $\lim_{n=\infty} c_n^{1/n} = 0$. The series

$$\varphi(z) = \sum_1^\infty \frac{c_n}{2^{2n}} z^{2n} \qquad (z = x + iy)$$

represents, therefore, an even entire function satisfying for large positive values of r the inequalities

$$4r^{-2} e^{C(\log \frac{1}{2} r)} \leq \sum_1^\infty \frac{c_n}{2^{2n}} r^{2n} = \varphi(r) = \max_{|z|=r} |\varphi(z)| \leq e^{C(\log r)}. \tag{8}$$

Consider the entire function

$$\Phi(z) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) e^{itz} dt.$$

For every $\alpha \,\varepsilon\, (0, 2\pi)$ we may write

$$|\Phi(z)| \leq \frac{1}{\pi} \int_{-\pi}^{\pi} |f(t)| e^{-ty} dt \leq \frac{1}{\pi} e^{-I(\alpha) + \pi|y|}$$

$$+ \frac{1}{\pi} \int_{-\pi+\alpha}^{\pi-\alpha} |f(t)| dt \cdot e^{(\pi-\alpha)|y|} < ce^{\pi|y|}(e^{-I(\alpha)} e^{-\alpha|y|}).$$

If then, $|y|$ being sufficiently large, we choose α such that $I(\alpha)/\alpha = |y|$, we have by the definition of the function ω

$$|\Phi(z)| < d_1 e^{\pi|y| - \omega(|y|)}. \tag{9}$$

It is clear that $\Phi(n) = a_n + ib_n$, $\Phi(-n) = a_n - ib_n$ $(n \geq 0)$.

Set $S_m(z) = \sum_1^m c_n z^{2n}$, where c_n are the Taylor coefficients of $\varphi(z)$, and consider the expression

$$Q(z) = \frac{F(z)\Phi(z)S_m\left(\frac{z-i}{2}\right)}{\sin \pi z} - \frac{1}{\pi}\sum_{n=-\infty}^{\infty} \frac{(-1)^n F(n)\Phi(n)S_m\left(\frac{n-i}{2}\right)}{z-n}. \tag{10}$$

Formulas and arguments of this type were considered by Levine and Lifschetz, but their reasoning is somewhat erroneous. A function such as $\varphi(z)$ was first introduced in this theory by the present author [3] in his proof of the generalization of de la Vallée-Poussin's theorem, but the construction of this function given here is different from that given previously. As a matter of fact the present function $\varphi(z)$ has special properties which we use in this paper.

From the relationships

$$\begin{aligned} &|F(n)\Phi(n)| < e^{-2C(\log|n|)} && (|n| \geq 1), \\ &\left|S_m\left(\frac{n-i}{2}\right)\right| < ke^{2m \log|n|} && (|n| \geq 1), \\ &\lim_{|n|=\infty} \frac{C(\log|n|)}{\log|n|} = \infty, \end{aligned} \tag{11}$$

it follows that

$$\sum_{n=-\infty}^{\infty} \left|F(n)\Phi(n)S_m\left(\frac{n-i}{2}\right)\right| < \infty,$$

and the series in (10) converges uniformly when z is outside the circles $|z - n| = \epsilon$ $(\epsilon > 0, n = 0, \pm 1, \pm 2, \cdots)$. We shall denote the region outside these circles by D_ϵ. Since the inequality $|(\sin \pi z)^{-1}| < ke^{-\pi|y|}$ is valid in D_ϵ, we have in D_ϵ

$$\left|\frac{F(z)\Phi(z)S_m\left(\frac{z-i}{2}\right)}{\sin \pi z}\right| < Ae^{\log M(r) - \omega(|y|) + 2m \log r}.$$

Denote by $D_{\epsilon,\gamma}$ that part of D_ϵ in which one of the two inequalities $|\arg z \pm \frac{1}{2}\pi| \leq \gamma < \frac{1}{2}\pi$ is satisfied. Since in $D_{\epsilon,\gamma}$ we have $|y| \geq r \cos \gamma$ $(r = |z|)$, we may write, for $r \geq r_0$,

$$\left|\frac{F(z)\Phi(z)S_m\left(\frac{z-i}{2}\right)}{\sin \pi z}\right| < Ae^{\log M(r) - (\cos \gamma)\omega(r) + 2m \log r}. \tag{12}$$

For, by elementary geometry we have for r sufficiently large: $(\cos \gamma)\omega(r) \leq \omega(r \cos \gamma)$ ($\omega(r) = I(\alpha)$ is non-decreasing to $+\infty$ as α tends to $+0$).

To every r sufficiently large there corresponds a quantity α, such that $r = I(\alpha)/\alpha$. We have $\omega(r) = I(\alpha)$, and thus

$$\frac{\log r}{\omega(r)} = \frac{\log I(\alpha) - \log \alpha}{I(\alpha)},$$

and from the second hypothesis (I) of the theorem, it follows that

$$\lim_{r=\infty} \frac{\log r}{\omega(r)} = 0.$$

It follows then from (12) and from the first hypothesis (I) of the theorem that in $D_{\epsilon,\gamma}$ the left member of (12) tends uniformly to zero when r tends to infinity. But since the series involved in (10) tends also uniformly to zero in $D_{\epsilon,\gamma}$ when r tends to infinity, we see that $Q(z)$ tends uniformly to zero in $D_{\epsilon,\gamma}$ when r tends to infinity. It is clear from the form of (10) that $Q(z)$ is an entire function. But, for r of the form $n + \frac{1}{2}$, where n is a positive integer, $| Q(z) | \leq LM(r)r^{2m}$. Therefore, for every r we have $| Q(z) | \leq LM(r + 1)(r + 1)^{2m}$.

From $\omega(r) = I(\alpha) = \alpha r$ and from the first hypothesis (I) of the theorem it follows that

$$\lim_{r=\infty} \frac{\log M(r)}{r} = \lim_{r=\infty} \frac{\alpha \log M(r)}{\omega(r)} = 0.$$

Therefore, for every $\epsilon > 0$ we have

$$| Q(z) | \leq e^{\epsilon r} \qquad (r > r_\epsilon). \tag{13}$$

Since on the straight lines $| \arg z \pm \frac{1}{2}\pi | = \gamma$, $Q(z)$ tends to zero when r tends to infinity, we see, by (13), on using Phragmén-Lindelöf's principle that in the part of the plane where one of the two inequalities $| \arg z \pm \frac{1}{2}\pi | \geq \gamma$ is satisfied, $Q(z)$ is bounded. But since $Q(z)$ tends uniformly to zero in $D_{\epsilon,\gamma}$, one sees that $Q(z)$ is identically zero. In other words, we have:

$$\frac{F(z)\Phi(z)S_m\left(\frac{z - i}{2}\right)}{\sin \pi z} = \frac{1}{\pi} \sum_{n=-\infty}^{\infty} \frac{(-1)^n F(n)\Phi(n)S_m\left(\frac{n - i}{2}\right)}{z - n}. \tag{14}$$

From the first inequality (11), from

$$\left|\varphi\left(\frac{n - i}{2}\right)\right| < e^{C(\log|n|)}, \qquad \left|S_m\left(\frac{n - i}{2}\right)\right| < e^{C(\log|n|)},$$

and from the fact that $S_m\left(\frac{z - i}{2}\right)$ tends uniformly to $\varphi\left(\frac{z - i}{2}\right)$ in every bounded closed region, it follows, on using (14), that in D_ϵ

$$\frac{F(z)\Phi(z)\varphi\left(\frac{z-i}{2}\right)}{\sin \pi z} = \frac{1}{\pi}\sum_{n=-\infty}^{\infty} \frac{(-1)^n F(n)\varphi\left(\frac{n-i}{2}\right)\Phi(n)}{z-n}. \tag{15}$$

Thus, the equality (15) holds throughout the plane.

It is evident from what precedes that

$$\frac{F(z)\Phi(z)}{\sin \pi z} = \frac{1}{\pi}\sum_{n=-\infty}^{\infty} \frac{(-1)^n F(n)\Phi(n)}{z-n}.$$

Since this sum is bounded in D_ϵ , one sees that

$$\Psi(z) = \frac{F(z+i)\Phi(z+i)}{\sin \pi(z+i)}$$

is bounded in the half plane $y \geq 0$. The series in (15) is also bounded in D_ϵ . Therefore, for x real, we have

$$\begin{aligned} | \Psi(x)\varphi(\tfrac{1}{2}x) | &= \left|\frac{F(x+i)\Phi(x+i)\varphi(\frac{1}{2}x)}{\sin \pi(x+i)}\right| \\ &= \frac{1}{\pi}\left|\sum_{n=-\infty}^{\infty} \frac{(-1)^n F(n)\Phi(n)\varphi\left(\frac{n-i}{2}\right)}{x+i-n}\right| < M < \infty. \end{aligned}$$

In other words,

$$0 < \varphi(\tfrac{1}{2}x) < M \mid \Psi(x) \mid^{-1} \qquad (\mid x \mid > 0). \tag{16}$$

But it is well known that, from the fact that $\Psi(z)$ is bounded for $y \geq 0$, and that $\Psi(z)$ is not identically zero, it follows that

$$\int_{-\infty}^{\infty} \frac{\log \mid \Psi(x) \mid dx}{x^2+1} > -\infty.$$

Thus, by (16),

$$\int_{-\infty}^{\infty} \frac{\log \varphi(\frac{1}{2}x)}{x^2+1}\, dx < \infty,$$

and from the first inequality (8) it follows at once, on setting $\log (r/2) = x$, that

$$\int^{\infty} C(x)e^{-x}dx < \infty.$$

Since $C(x)$ is an arbitrary convex function not greater than $\frac{1}{2}\Pi(x)$, we have the desired inequality:

$$\int^{\infty} \Pi(x)e^{-x}dx < \infty.$$

The proof of the theorem is complete.

Applications of the principal theorem. At first we shall apply the principal theorem by choosing for $F(z)$ different entire functions; we shall then obtain various types of statements concerning quasi-analyticity.

Suppose that in $[-\pi, \pi]$, $f(x) \in C\{M_n\}$, with $f^{(n)}(-\pi) = f^{(n)}(\pi) = 0$ $(n \geq 0)$, where $f(x)$ is not identically zero. By Taylor's formula one sees that for $x \in (-\pi, \pi)$

$$f(x) = \frac{f^{(n)}(x_1)}{n!}(x + \pi)^n = \frac{f^{(n)}(x_2)}{n!}(x - \pi)^n \qquad (-\pi < x_1 < x_2 < \pi).$$

The second condition (I) of the principal theorem is an immediate consequence.

If we choose $F(z) \equiv 1$, the first condition (I) is also satisfied. In addition we have

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) \cos nx \, dx = \frac{\pm 1}{\pi n^p}\int_{-\pi}^{\pi} f^{(p)}(x) \genfrac{}{}{0pt}{}{\cos nx}{\sin nx} \, dx \qquad (n \geq 1).$$

Thus

$$| a_n | \leq \frac{2k^p M_p}{n^p} \qquad (p \geq 1, n \geq 1),$$

and

$$| a_n | \leq \frac{2}{T(n/k)} \qquad (n \geq 1),$$

where

$$T(r) = \underset{1 \leq p}{\text{l.u.b.}} \frac{r^p}{M_p}.$$

Similarly,

$$| b_n | \leq \frac{2}{T(n/k)} \qquad (n \geq 1).$$

In other words,

$$-\log | a_n + ib_n | \geq -\log 4 + \log T\left(\frac{n}{k}\right) \qquad (n \geq 1).$$

But $\log T(e^x)$ is a convex function in the smallest interval I containing the projection E_x of all those points of the set $\{(\log n, -\log | a_n + ib_n |)\}$ for which $| a_n + ib_n | > 0$. Therefore, if $\Pi(x)$ denotes the corresponding base, we have in I, $-\log 4 + \log T(e^x/k) \leq \Pi(x)$. For other values of x, $\Pi(x) = \infty$. It follows from the principal theorem that

$$\int^{\infty} \log T(e^x)e^{-x}dx < \infty,$$

a relation equivalent to

$$\int^{\infty} \frac{\log T(r)}{r^2}\, dr < \infty .$$

If we refer to the introduction, we see that the fundamental theorem of Denjoy-Carleman (the sufficient condition for quasi-analyticity), in the form B, is a particular case of our main theorem.

Suppose that $f(x)$ is integrable in $(-\pi, \pi)$, periodic with period 2π, different from zero in a set of positive measure, and equal to zero in an interval $(-\pi - \alpha_1, -\pi + \alpha_1)$. For this function $I(\alpha) = \infty$ for $\alpha \leq \alpha_1$. Therefore, the second condition (I) of the theorem is satisfied. It is evident that in this case

$$\liminf_{r=\infty} \frac{\omega(r)}{r} \geq \alpha_1 .$$

Let now $\{n_i\}$ be a sequence of increasing positive integers such that $\lim_{i=\infty} n_i/i = \infty$. If we set

$$F(z) = \prod_{i=1}^{\infty} \left(1 - \frac{z^2}{n_i^2}\right),$$

we see, by the well-known facts of the theory of entire functions, that

$$\lim_{r=\infty} \frac{\log M(r)}{r} = 0,$$

and therefore that

$$\lim_{r=\infty} \frac{\log M(r)}{\omega(r)} = 0.$$

Thus, the first condition (I) of the theorem is satisfied.

But $F(n_i) = 0$ $(i \geq 1)$, and the base $\Pi(x)$ of the sequence $\{(\log n, -\log |(a_n + ib_n)F(n)|)\}$ is that of the sequence $\{(\log m_i, -\log |(a_{m_i} + ib_{m_i})F(m_i)|)\}$, where $\{m_i\}$ is the sequence of positive integers different from n_i $(i \geq 1)$. Therefore, by the theorem, we see that if $f(x)$ is integrable in $(-\pi, \pi)$, if this function is different from zero in a set of positive measure, but is zero in an interval of positive length, and if, for a sequence $\{n_i\}$, with $\lim n_i/i = \infty$, we place $F(z) = \prod_{i=1}^{\infty} \left(1 - \frac{z^2}{n_i^2}\right)$, then, on denoting by $\{m_i\}$ the set of positive integers different from the n_i $(i \geq 1)$, the base $\Pi(x)$ of the sequence $\{(\log m_i, -\log |(a_{m_i} + ib_{m_i}) \times F(m_i)|)\}$ is such that

$$\int^{\infty} \Pi(x) e^{-x} dx < \infty .$$

By a simple transformation it is seen that in these statements the point $x = -\pi$ may be replaced by any other point $x = x_0$. This transformation $x = t - \pi$ conserves the values $| a_n + ib_n |$.

In particular, the coefficients a_{m_i}, b_{m_i} cannot all be zero; in other words, if the function $f(x)$, integrable in $(-\pi, \pi)$, is zero in an interval of positive length (for instance in an interval around $-\pi$), and if

$$f(x) \sim \sum (a_{n_i} \cos n_i x + b_{n_i} \sin n_i x),$$

with $\lim n_i/i = \infty$, then $f(x)$ is zero almost everywhere.

This last particular statement follows also easily from the known facts on Taylor series admitting the circle of convergence as a cut, and Schwartz's reflexion principle [3].

Suppose now that $f(x)$ is integrable, periodic with period 2π, different from zero in a set of positive measure, and such that

$$\liminf_{\alpha = +0} \frac{\log I(\alpha)}{-\log \alpha} = \rho > 0.$$

Then obviously the second condition (I) of our principal theorem is satisfied; but, if $\epsilon > 0$ is smaller than ρ, from $I(\alpha) > \alpha^{-\rho+\epsilon}$ $(\alpha < \alpha_\epsilon)$, on setting $\alpha^{-\rho+\epsilon} = \omega_1(\alpha^{-\rho-1+\epsilon})$, we have $\omega(r) \geq \omega_1(r)$ for r sufficiently large. Since $\omega_1(r) = r^{(\rho-\epsilon)/(1+\rho-\epsilon)}$, we see that

$$\omega(r) > r^{(\rho-\epsilon)/(1+\rho-\epsilon)} \qquad (r > r_\epsilon).$$

Let $\{n_i\}$ be a sequence of positive integers of which the exponent of convergence is $\sigma < 1$, and let

$$F(z) = \prod_{i=1}^{\infty} \left(1 - \frac{z^2}{n_i^2}\right).$$

Clearly

$$\log M(r) < r^{\sigma+\epsilon} \qquad (r > r'_\epsilon).$$

Therefore, if

$$\rho > \frac{\sigma}{1 - \sigma},$$

it will follow that

$$\lim_{r=\infty} \frac{\log M(r)}{\omega(r)} = 0.$$

For these two functions $f(x)$, $F(z)$, the hypotheses of our theorem are satisfied. As above, the base of the sequence $\{(\log n, -\log | (a_n + ib_n)F(n) |)\}$ is that of

the sequence $\{(\log m_i, -\log |(a_{m_i} + ib_{m_i})F(m_i)|)\}$, where $\{m_i\}$ is the sequence of positive integers different from all the n_i $(i \geq 1)$.

We may therefore state the following result.

If $f(x)$ is integrable in $(-\pi, \pi)$, is different from zero in a set of positive measure, is periodic with period 2π, and satisfies

$$\liminf_{\alpha \to +0} \frac{I(\alpha)}{-\log \alpha} = \rho > 0,$$

then, if we denote by $\{n_i\}$ any sequence of positive integers with the exponent of convergence σ satisfying the inequality

$$\rho > \frac{\sigma}{1 - \sigma},$$

and by $\{m_i\}$ the sequence of positive integers different from the n_i $(i \geq 1)$, we have, on denoting by $F(z)$ the entire function $\prod_{i=1}^{\infty} \left(1 - \frac{z^2}{n_i^2}\right)$, and by $\Pi(x)$ the base of the sequence $\{(\log m_i, -\log |(a_{m_i} + ib_{m_i})F(m_i)|)\}$,

$$\int^{\infty} \Pi(x)e^{-x}dx < \infty.$$

This may be reduced to a theorem of form E.

Theorem C mentioned in the introduction, with the difference that in the definition of an exponential zero "lim sup" has to be replaced by "lim inf", and the expressions "right-hand mean-value zero of exponential order ρ" or "left-hand $\cdots$" have to be replaced by "mean-value zero of exponential order ρ on both sides", also follows as a particular case of the last statement.

The other theorems mentioned in the introduction, with similar modifications, are also particular cases of the principal theorem.

We shall now choose $f(x)$ in a special manner and get a statement on entire functions.

Let $F(z)$ be any even entire function such that

$$\lim_{r=\infty} \frac{\log M(r)}{r} = 0. \tag{17}$$

Let $0 < \beta < 1$, and consider the function $f(x)$ defined as unity in the interval $[-\beta\pi, \beta\pi]$, and zero in the rest of the interval $[-\pi, \pi]$. Then $|a_n + ib_n| = 2\pi^{-1}n^{-1} |\sin \beta n\pi|$ $(n \geq 1)$. The second condition (I) of the theorem is satisfied, and $\lim_{r=\infty} \omega(r)/r = (1 - \beta)\pi$. Hence, the first condition (I) of the theorem is also satisfied.

Thus, for an even entire function satisfying (17) the base $\Pi(x)$ of the sequence $\{(\log n, -\log|\sin(\beta n\pi)n^{-1}F(n)|)\}$ satisfies the inequality

$$\int^{\infty} \Pi(x)e^{-x}dx < \infty.$$

Clearly, the same inequality is valid for the base of the sequence

$$\{(\log n, -\log|F(n)|)\}.$$

This statement, like the main theorem itself, expresses the fact that an entire function of which the maximum $M(r)$ does not increase too rapidly has a certain character of flatness on every straight line.

BIBLIOGRAPHY

1. T. CARLEMAN, *Les fonctions quasi analytiques*, Paris, Gauthier-Villars, 1926.
2. B. LEVINE AND M. LIFSCHETZ, *Quasi-analytic functions represented by Fourier series*, Recueil Mathématique, Moscow, vol. 9(51), n. 3(1941).
3. S. MANDELBROJT, *Séries de Fourier et classes quasi-analytiques de fonctions*, Paris, Gauthier-Villars, 1935.
4. S. MANDELBROJT ET N. WIENER, Two notes, Comptes Rendus de l'Académie des Sciences, Paris, vol. 203(1936).
5. CH. DE LA VALLÉE-POUSSIN, The Rice Institute Pamphlet, vol. xvii(1925).
6. S. MANDELBROJT, *Analytic Functions and Classes of Infinitely Differentiable Functions*, The Rice Institute Pamphlet, vol. xxix(1942).

THE RICE INSTITUTE.

FUNCTIONS OF BOUNDED TYPE

By Audrey Wishard

Introduction. By definition, a function $f(z)$ is of bounded type (beschränktartig [5]) in a region G, if

$$(1) \qquad f(z) = W_2(z)/W_1(z) \qquad \text{in } G,$$

where $|W_i(z)| \leq 1$ and $W_i(z)$ are analytic interior to G $(i = 1, 2)$.

A given meromorphic function $f(z)$ is of bounded type in a simply connected region G if and only if there exists a function $U_1(z)$, positive and harmonic in G and such that

$$(2) \qquad \log |f(z)| \leq U_1(z) \qquad \text{in } G.$$

(Throughout this paper we shall use "harmonic" to describe functions which are harmonic in the strict sense except at isolated points b, where they behave like $\pm k \log |z - b|$.) For, if (1) holds, then so must (2) with $U_1 = \log |1/W_1|$, whereas, if (2) holds, then $\log |f(z)| = U_1(z) - U_2(z)$, where $U_2(z)$ is also positive and harmonic in G. The fact that one can obtain from $U_i(z)$ analytic functions $W_i(z)$ such that $U_i = \log |1/W_i|$ is discussed by R. Nevanlinna [5].

Our problem is to determine under what conditions $f(z)$, meromorphic in the interior of a region G, and perhaps on part of its boundary, is of bounded type in G, and, if $f(z)$ is of bounded type, to find a representation for $\log |f(z)|$. (The reader will notice that in the theorems of this paper $\log |f(z)|$ may be replaced by a harmonic function $U(z)$. The theorems then give conditions under which $U(z)$ may be expressed as the difference between two positive harmonic functions in G.) Explicit solutions of this problem have been given in two instances by R. Nevanlinna. The first solution [4] is for the case in which G is the halfplane, $x > 0$, $f(z)$ being meromorphic when $x \geq 0$; the second [5] is for the case in which G is the unit circle and $f(z)$ is meromorphic interior to G. The case in which G is a strip, $a < x < b$, has been considered by E. Hille [3] in a closely related problem, and by L. V. Ahlfors [1] in a theorem the proof of which demands more restrictive conditions on $f(z)$ than those which we shall use. The conditions found by these writers are of two kinds: (1) a certain sum extended over the poles of $f(z)$ in G must be convergent; (2) a mean value of the function, expressed as a weighted integral, on curves which approach the boundary of G, must remain finite. The reader will find conditions of both types in the theorems below.

Received July 30, 1940; in revised form, September 29, 1942; presented to the American Mathematical Society, December, 1937. This paper is based on the material of the author's doctoral dissertation presented at Radcliffe College in 1938. The work was done under the direction of Professor L. V. Ahlfors, to whom the author wishes to express her great indebtedness.

1. We give first a brief account of the basic tools which are to be used in this paper. For further details and explanation concerning the methods and terminology, the reader is referred to the writings of R. Nevanlinna which have been mentioned above.

The first two theorems which we shall prove give conditions that $f(z)$ shall be of bounded type in a region: $0 < x < \pi$. The following notations will be used: $G(X, Y)$ is the region: $X < x < \pi - X$, $-Y < y < Y$ $(0 < X < \frac{1}{2}\pi, 0 < Y)$, $\Gamma(X, Y)$ is its boundary and $g(\zeta, z; X, Y)$ is its Green's function, with positive pole at $\zeta = z$. We number the four lines which form the boundary in counterclockwise order, starting with the left vertical side. Then $\Gamma(X, Y) = \Gamma_1(X, Y) + \cdots + \Gamma_4(X, Y)$. We write $G(0, Y)$, $\Gamma_j(0, Y)$, and $g(\zeta, z; 0, Y)$ more simply as $G(Y)$, $\Gamma_j(Y)$, and $g(\zeta, z; Y)$. These, in turn, we write as G, Γ_j, and $g(\zeta, z)$ when $Y = \infty$ (where $j = 1, 3$). As usual, $\log^+ |f| = \max [\log |f|, 0]$. The poles of $f(z)$ are denoted by b_ν $(b_\nu = \beta_\nu + i\gamma_\nu)$ and the sum: $\sum_G g(z, b_\nu)$ is to be taken over all the poles of $f(z)$ in G.

It is well known that $g(\zeta, z; X, Y)$ is a continuous function of all of its arguments and that, if $G(X, Y) \subset G(X', Y')$, then

$$g(\zeta, z; X, Y) \leq g(\zeta, z; X', Y'), \qquad X' \leq X, \qquad Y' \geq Y. \tag{3}$$

The derivative $\partial g(\zeta, z; X, Y)/\partial n$ is taken with respect to ζ at a point on $\Gamma(X, Y)$ in the direction of the inner normal. The interpretation of this derivative at the vertices of $G(X, Y)$ is of no significance. It is known that $\partial g(\zeta, z; X, Y)/\partial n$ is continuous in all four arguments.

Choose X and Y so that $f(z)$ has no poles or zeros on $\Gamma(X, Y)$. Then $f(z)$ is meromorphic on $G(X, Y) + \Gamma(X, Y)$ and is continuous on $\Gamma(X, Y)$. Let $U_1(z; X, Y)$ be the harmonic function with boundary values $\log^+ |f(z)|$ on $\Gamma(X, Y)$ and with logarithmic poles at the poles b_ν of $f(z)$ in $G(X, Y)$. Then (Green's theorem, cf. [2])

$$U_1(z; X, Y) = \int_{\Gamma(X,Y)} \log^+ |f(\zeta)| \frac{\partial}{\partial n} g(\zeta, z; X, Y) \, |d\zeta| + \sum_{G(X,Y)} g(z, b_\nu; X, Y). \tag{4}$$

Since $\log |f(z)| - U_1(z; X, Y)$ is harmonic and has no positive poles in $G(X, Y)$, and since its boundary values are non-positive and continuous on $\Gamma(X, Y)$, we have $\log |f(z)| \leq U_1(z; X, Y)$ in $G(X, Y)$. But a simple argument shows that $U_1(z; X, Y)$ is a continuous function of X and Y, for z fixed interior to $G(X, Y)$, even when $f(z)$ has poles on $\Gamma(X, Y)$. Hence

$$\log |f(z)| \leq U_1(z; X, Y) \qquad (0 < X < \tfrac{1}{2}\pi, 0 < Y < \infty).$$

A now classical argument in the Nevanlinna theory shows: (1) that $U_1(z; X, Y)$ is the smallest positive harmonic function in $G(X, Y)$ which $\geq \log |f(z)|$ (that

is, if $U(z) \geq 0$ and is harmonic in $G(X, Y)$ and if $\log | f(z) | \leq U(z)$ in $G(X, Y)$, then $U_1(z; X, Y) \leq U(z)$ for every z in $G(X, Y)$); (2) that $\lim_{G(X,Y) \to G} U_1(z; X, Y)$ exists and is either a positive harmonic function in G or is infinite for every z in G; (3) that *a necessary and sufficient condition that $f(z)$ be of bounded type in G is that* $\lim_{G(X,Y)\to G} U_1(z; X, Y)$ *be a harmonic function for $z \subset G$.*

2. We first prove the following theorem, using the notation of §1.

THEOREM A. *Suppose that $f(z)$ is meromorphic in G and on Γ. Then $f(z)$ is of bounded type in G if and only if the following three conditions hold:*

$$\sum_G e^{-|\gamma_\nu|} \sin \beta_\nu < \infty, \tag{I}$$

$$\int_{-\infty}^{\infty} \log^+ | f(x + iy) | e^{-|y|} \, dy < \infty \qquad \textit{for } x = 0 \textit{ and } x = \pi, \tag{II}$$

$$\varliminf_{y\to\infty} e^{-y} [m(y) + m(-y)] < \infty, \quad \textit{where } m(y) = \int_0^\pi \log^+ | f(x + iy) | \sin x \, dx. \tag{III}$$

If $f(z)$ is of bounded type in G, then $\log | f(z) | \leq U_1(z)$ *in G, where*

$$U_1(z) = \int_\Gamma \log^+ | f(\zeta) | \frac{\partial}{\partial n} g(\zeta, z) | d\zeta | + \sin x [\eta_1 e^y + \eta_2 e^{-y}] + \sum_G g(z, b_\nu) \tag{5}$$

and

$$\eta_1 = \varliminf_{y\to\infty} e^{-y} m(y)$$

and

$$\eta_2 = \varliminf_{y\to\infty} e^{-y} m(-y).$$

Since $f(z)$ is meromorphic on $x = 0$ and on $x = \pi$, the remarks of §1 apply even when $X = 0$. By definition, $U_1(z; Y) = U_1(z; 0, Y)$. We investigate $\lim_{Y\to\infty} U_1(z; Y)$. It will be shown that this limit is $U_1(z)$, as given by (5), and that $U_1(z)$ is finite if and only if (I), (II), and (III) hold.

First, we prove that, if either of the indicated limits is finite,

$$\lim_{G(X,Y)\to G} \sum_{G(X,Y)} g(z, b_\nu ; X, Y) = \sum_G g(z, b_\nu). \tag{6}$$

By using (3), it is easy to prove that $\lim_{G(X,Y)\to G} g(z, b; X, Y) = g(z, b)$ and that, when $G(X', Y') \subset G(X, Y)$, we have

$$\sum_{G(X',Y')} g(z, b_\nu; X, Y) \le \sum_{G(X,Y)} g(z, b_\nu; X, Y) \le \sum_G g(z, b_\nu).$$

Then

$$\sum_{G(X',Y')} g(z, b_\nu) \le \overline{\lim_{G(X,Y)\to G}} \sum_{G(X,Y)} g(z, b_\nu; X, Y) \le \sum_G g(z, b_\nu).$$

Letting $G(X', Y')$ approach G, we get (6).

Now $\sum_G g(z, b_\nu)$ converges to a harmonic function if and only if

$$\sum_{\substack{G \\ b_\nu \ne \frac{1}{2}\pi}} g(\tfrac{1}{2}\pi, b_\nu)$$

converges. If $z = x + iy$, $\zeta = s + it$, $b = \beta + i\gamma$, and z^* denotes the conjugate of z,

$$g(\zeta, z) = \log \left| \frac{e^{i\zeta} - e^{-iz^*}}{e^{i\zeta} - e^{iz}} \right|,$$

$$g(\zeta, z) = \tfrac{1}{2} \log \frac{e^{y-t} + e^{t-y} - 2\cos(s + x)}{e^{y-t} + e^{t-y} - 2\cos(s - x)}, \tag{7}$$

$$g(\tfrac{1}{2}\pi, b) = \tfrac{1}{2} \log \left[1 + \frac{4 \sin \beta}{e^\gamma + e^{-\gamma} - 2 \sin \beta} \right].$$

From this last equation it is easy to see that $\sum_G g(z, b_\nu)$ converges if and only if (I) is satisfied.

The integral in the expression (4) for $U_1(z; Y)$ may be expressed as the sum of four integrals:

$$A_j(z; Y) = \int_{\Gamma_j(Y)} \log^+ | f(\zeta) | \frac{\partial}{\partial n} g(\zeta, z; Y) \, | d\zeta | \qquad (j = 1, 2, 3, 4).$$

For $j = 1, 3$ the integrand is a non-negative increasing function of Y. Hence, whenever either of the limits is finite,

$$A_j(z) = \lim_{Y\to\infty} A_j(z, Y) = \int_{\Gamma_j} \log^+ | f(\zeta) | \frac{\partial}{\partial n} g(\zeta, z) \, | d\zeta | \qquad (j = 1, 3).$$

Now $A_j(z)$ converges to a harmonic function in G if and only if $A_j(\frac{1}{2}\pi)$ converges. This gives us (II), for

$$\frac{\partial}{\partial s} g(\zeta, \tfrac{1}{2}\pi) \Big|_{s=0} = \frac{2}{e^t + e^{-t}} = \frac{\partial}{\partial n} g(\zeta, \tfrac{1}{2}\pi) \Big|_{\Gamma_j} \qquad (j = 1, 3).$$

The terms $A_2(z; Y)$ and $A_4(z; Y)$ remain to be considered. We shall make use of the fact that, if the first two conditions of our theorem are fulfilled, then $\lim_{Y\to\infty} [A_2(z; Y) + A_4(z; Y)]$ exists, finite or infinite. This is easily seen to be true by making use of the expression

$$[A_2 + A_4(z; Y)] = U_1(z; Y) - [A_1 + A_3 + \sum_{G(Y)} g(z, b_\nu ; Y)]$$

because $U_1(z; Y)$ is a non-decreasing function of Y.

We wish to approximate $\frac{\partial}{\partial n} g(\zeta, z; Y)$ on $\Gamma_4(Y)$. We notice that the points z and $z_1 = z^* + 2iY$ are symmetric with respect to $\Gamma_4(Y)$. Therefore, $g(\zeta, z) = g(\zeta, z_1)$ for ζ on $\Gamma_4(Y)$. Using the principle of the maximum one can easily prove that

$$F_1(\zeta) = g(\zeta, z) - g(\zeta, z_1) - g(\zeta, z; Y) = 0$$

for ζ on $\Gamma_4(Y)$ and that $F_1(\zeta) \geq 0$ for ζ in $G(Y)$. Therefore,

$$\frac{\partial}{\partial n} F_1(\zeta) \bigg|_{\Gamma_4(Y)} \geq 0,$$

and

$$\frac{\partial}{\partial n} g(\zeta, z; Y)\bigg|_{\Gamma_4(Y)} \leq \frac{\partial}{\partial n} [g(\zeta, z) - g(\zeta, z_1)]_{\Gamma_4(Y)} = 2 \frac{\partial}{\partial n} g(\zeta, z)\bigg|_{\Gamma_4(Y)}.$$

Similarly, $z_2 = z + 4iY$ and $z_3 = z^* - 2iY$ are symmetric in $\Gamma_4(Y)$ and we can prove that, if

$$F_2(\zeta) = g(\zeta, z) - g(\zeta, z_1) + g(\zeta, z_2) - g(\zeta, z_3) - g(\zeta, z; Y),$$

then $\frac{\partial}{\partial n} F_2(\zeta) \leq 0$ on $\Gamma_4(Y)$. Therefore,

$$\frac{\partial}{\partial n} g(\zeta, z; Y) \,|_{\Gamma_4(Y)} \geq 2 \frac{\partial}{\partial n} [g(\zeta, z) - g(\zeta, z_3)]_{\Gamma_4(Y)} \,.$$

Again, differentiating (7), we find that

$$\frac{\partial}{\partial n} g(\zeta, z) \,|_{\Gamma_4(Y)} = 2 \frac{[e^{Y-y} - e^{y-Y}] \sin s \sin x}{[e^{Y-y} + e^{y-Y} - 2 \cos (s + x)][e^{Y-y} + e^{y-Y} - 2 \cos (s - x)]},$$

$$\frac{\partial}{\partial n} g(\zeta, z) \,|_{\Gamma_4(Y)} \leq 2e^{y} \sin x \, e^{-Y} \sin s \frac{e^{2(Y-y)}}{[e^{Y-y} - 2]^2},$$

$$\frac{\partial}{\partial n} g(\zeta, z) \,|_{\Gamma_4(Y)} \geq 2e^{y} \sin x \, e^{-Y} \sin s \frac{e^{2(Y-y)} - 1}{[e^{Y-y} + 1]^2},$$

$$-\frac{\partial}{\partial n} g(\zeta, z_3) \,|_{\Gamma_4(Y)} \geq -2e^{y} \sin x \, e^{-Y} \sin s \frac{e^{4Y}}{[e^{3Y+y} - 2]^2}.$$

Therefore, $\frac{\partial}{\partial n} g(\zeta, z; Y) |_{\Gamma_4(Y)} = 4e^y \sin x\ e^{-Y} \sin s\ \phi_1(y, Y)$, where $\lim_{Y\to\infty} \phi_1(y, Y) = 1$ for y fixed. Then

$$(8) \qquad A_4(z; Y) = \frac{2}{\pi} e^y \sin x\ \phi_1(y, Y) e^{-Y} m(Y).$$

Similarly,

$$(9) \qquad A_2(z; Y) = \frac{2}{\pi} e^{-y} \sin x\ \phi_2(y, Y) e^{-Y} m(-Y),$$

where $\lim_{Y\to\infty} \phi_2(y, Y) = 1$ for y fixed.

Since $\lim_{Y\to\infty} [A_2(z; Y) + A_4(z; Y)]$ exists, finite or infinite, assuming that (I) and (II) hold, the limit is finite if and only if (III) also holds.

Suppose, finally, that $\lim_{Y\to\infty} [A_2(z; Y) + A_4(z; Y)]$ is finite. Then the following limits are finite:

$$\eta_1 = \underline{\lim}_{Y\to\infty} e^{-Y} m(Y), \qquad \eta_1' = \overline{\lim}_{Y\to\infty} e^{-Y} m(Y),$$

$$\eta_2 = \underline{\lim}_{Y\to\infty} e^{-Y} m(-Y), \qquad \eta_2' = \overline{\lim}_{Y\to\infty} e^{-Y} m(-Y).$$

One can prove by elementary methods that

$$\begin{aligned}\lim_{Y\to\infty} [A_2(z; Y) + A_4(z; Y)] &= \underline{\lim}_{Y\to\infty} A_2(z; Y) + \overline{\lim}_{Y\to\infty} A_4(z; Y)\\ &= \overline{\lim}_{Y\to\infty} A_2(z; Y) + \underline{\lim}_{Y\to\infty} A_4(z; Y).\end{aligned}$$

This gives

$$\frac{2}{\pi} \sin x[\eta_1 e^y + \eta_2' e^{-y}] = \frac{2}{\pi} \sin x[\eta_1' e^y + \eta_2 e^{-y}].$$

This is impossible unless $\eta_1 = \eta_1'$ and $\eta_2' = \eta_2$, whence

$$\lim_{Y\to\infty} A_4(z; Y) = \frac{2}{\pi} \eta_1 e^y \sin x, \qquad \lim_{Y\to\infty} A_2(z; Y) = \frac{2}{\pi} \eta_2 e^{-y} \sin x.$$

3. THEOREM B. *Suppose that $f(z)$ is meromorphic in the interior of G. Then $f(z)$ is of bounded type in G if and only if the following two conditions are satisfied:*

$$(\mathrm{I}) \qquad \sum_G e^{-|\gamma_\nu|} \sin \beta_\nu < \infty,$$

(IV) *constants C_1 and C_2 exist such that, when $0 < x < \pi$,*

$$\int_{-y}^{y} \log^+ | f(x + it) | e^{-|t|} dt \leq C_1 + C_2 y \sin x.$$

We return to our expression (4) for $U_1(z; X, Y)$. As in Theorem A, (I) is the condition that $\lim_{G(X,Y)\to G} \sum_{G(X,Y)} g(z, b_\nu; X, Y)$ be finite. Again we write the integral in the expression for $U_1(z; X, Y)$ as the sum of four integrals $A_j(z; X, Y)$.

Assuming first that (IV) holds, we wish to show that $\lim_{X\to 0} A_j(\frac{1}{2}\pi; X, Y)$ is bounded for all $Y > 0$. Now the region $G(X, Y)$ is contained in the region: $X < x < X + \pi, -\infty < y < \infty$ and the two regions have $\Gamma_1(X, Y)$ as a common boundary line. The Green's function of the second region, found by mapping the region on G, is $g(\zeta - X, z - X)$. Therefore,

$$\frac{\partial}{\partial n} g(\zeta, \tfrac{1}{2}\pi; X, Y) \,|_{\Gamma_1(X,Y)} \leq \frac{\partial}{\partial n} g(\zeta - X, \tfrac{1}{2}\pi - X) \,|_{\Gamma_1(X,Y)}$$

$$= 2 \frac{\sin (\frac{1}{2}\pi - X)}{e^t + e^{-t} - 2 \cos (\frac{1}{2}\pi - X)}.$$

When X is sufficiently small ($X < \frac{1}{4}$),

$$A_1(\tfrac{1}{2}\pi; X, Y) \leq \frac{2}{\pi} \int_{-Y}^{Y} \log^+ | f(X + it) | \, e^{-|t|} dt \leq C_1 + C_2 Y \sin X.$$

The same inequality must hold for $A_3(\frac{1}{2}\pi; X, Y)$. As before, $U_1(\frac{1}{2}\pi; Y)$ will be bounded if $\overline{\lim}_{y\to\infty} e^{-y} [m(y) + m(-y)]$ is finite. But

$$\begin{aligned} \int_{-y}^{y} e^{-|t|} m(t) \, dt &= \int_{-y}^{y} e^{-|t|} \int_0^{\pi} \log^+ | f(x + it) | \sin x \, dx \, dt \\ &= \int_0^{\pi} \sin x \int_{-y}^{y} \log^+ | f(x + it) | \, e^{-|t|} \, dt \, dx \qquad (10) \\ &\leq \int_0^{\pi} (C_1 + C_2 y \sin x) \, dx = \pi C_1 + 2C_2 y. \end{aligned}$$

This cannot be true unless $\overline{\lim}_{y\to\infty} e^{-y} [m(y) + m(-y)]$ is finite. Thus we have proved that if (I) and (IV) hold, $f(z)$ is of bounded type.

Suppose that $f(z)$ is of bounded type. Let $U(z)$ be a positive harmonic function in G such that $\log | f(z) | \leq U(z)$ in G. We define

$$M(x, y) = \int_{-y}^{y} [U(x + it) + U(\pi - x + it)][e^{-|t|} - e^{|t|-2y}] \, dt$$

and compare $M(x, y)$ with the integral which appears in condition IV. When $y > \log 2$,

$$M(x, y) \geq \int_{-y+\log 2}^{y-\log 2} [U(x + it) + U(\pi - x + it)][e^{-|t|} - e^{-|t|-2\log 2}] \, dt$$

$$\geq \tfrac{3}{4} \int_{-y'}^{y'} [\log^+ | f(x + it) | + \log^+ | f(\pi - x + it) |] \, e^{-|t|} \, dt,$$

where $y' = y - \log 2$. We shall prove that (IV) is satisfied by showing that constants C_1' and C_2' exist such that $M(x, y) \leq C_1' + C_2' y \sin x$.

The function $\lambda(x, y) = [e^{-|y|} - e^{|y|-2Y}] \sin x$ is harmonic in G except on the line $y = 0$, where it is continuous. Let X and Y be chosen so that $U(z)$ has no poles on $\Gamma(X, Y)$. Following the standard method, we remove from $G(X, Y)$ circles of radius r about the poles of $U(z)$ and we remove the line $y = 0$. Applying Green's theorem in the region thus formed, and letting r become zero, we get:

$$\int_{\Gamma(X,Y)} [\lambda \frac{\partial}{\partial n} U - U \frac{\partial}{\partial n} \lambda] \mid dz \mid - 2 \int_X^{\pi - X} U(x) \frac{\partial}{\partial y} \lambda \mid_{y=0+} dx$$

$$- 2\pi \sum_{G(X,Y)} \lambda(\beta_\nu , \gamma_\nu) = 0.$$

Since $U(z)$ is positive and harmonic in G, we may apply to $U(z)$ the results of the first part of this paper. If we form the functions $A_j(z; X, Y)$, using $U(z)$ instead of $\log | f(z) |$, we see that $A_j(z; X, Y)$ must be bounded for z fixed. Using (8) and (9), we see that $e^{-|Y|} \int_0^\pi U(x + iY) \sin x \, dx$ must be bounded for all Y. Thus the following quantities are bounded for all $0 < X < \frac{1}{2}\pi$, $0 < Y$:

$$\int_{\Gamma_j(X,Y)} \left[\lambda \frac{\partial}{\partial n} U - U \frac{\partial}{\partial n} \lambda\right] \mid dz \mid \tag{11}$$

$$= 2e^{-Y} \int_X^{\pi - X} U(x + iy) \mid_{\Gamma_j(X,Y)} \sin x \, dx \qquad (j = 2, 4),$$

$$\sum_{G(X,Y)} \lambda(\beta_\nu , \gamma_\nu) = \sum_{G(X,Y)} [e^{-|\gamma_\nu|} - e^{|\gamma_\nu|-2Y}] \sin \beta_\nu ,$$

$$- 2 \int_X^{\pi - X} U(x) \frac{\partial}{\partial y} \lambda \mid_{y=0+} dx = 2[1 + e^{-2Y}] \int_X^{\pi - X} U(x) \sin x \, dx.$$

Thus we see that positive constants K_1 and K_2 exist such that

$$-K_1 \leq \int_{\Gamma_1(X,Y)+\Gamma_3(X,Y)} \left[\lambda \frac{\partial}{\partial n} U - U \frac{\partial}{\partial n} \lambda\right] dz$$

$$= \sin X \frac{\partial}{\partial X} M(X, Y) - \cos X \, M(X, Y) \leq K_2 ,$$

$$-K_1 \leq \sin^2 x \frac{\partial}{\partial x} \frac{M(x, y)}{\sin x} \leq K_2 ,$$

$$K_1 \frac{\partial}{\partial x} \cot x \leq \frac{\partial}{\partial x} \frac{M(x, y)}{\sin x} \leq -K_2 \frac{\partial}{\partial x} \cot x.$$

Although $\partial M(x, y)/\partial x$ may have jumps, $M(x, y)$ itself is continuous. Therefore, as x increases from 0 to $\frac{1}{2}\pi$,

$$\frac{M(x, y)}{\sin x} - K_1 \cot x \qquad \text{is increasing,}$$

$$\frac{M(x, y)}{\sin x} + K_2 \cot x \qquad \text{is decreasing,}$$

$$\begin{aligned} &M(x, y) \leq K_1 \cos x + M(\tfrac{1}{2}\pi, y) \sin x, \\ &M(x, y) + K_2 \cos x \geq M(\tfrac{1}{2}\pi, y) \sin x. \end{aligned} \tag{12}$$

Multiply the second of the above inequalities by $\sin x$ and integrate from 0 to $\frac{1}{2}\pi$. We get

$$M(\tfrac{1}{2}\pi, y) \int_0^{\frac{1}{2}\pi} \sin^2 x \, dx \leq \tfrac{1}{2}K_2 + \int_0^{\frac{1}{2}\pi} M(x, y) \sin x \, dx.$$

But

$$\int_0^{\frac{1}{2}\pi} M(x, y) \sin x \, dx = \int_{-y}^{y} [e^{-|t|} - e^{|t|-2y}] \int_0^{\pi} U(x + it) \sin x \, dx \, dt.$$

Using equation (11), this gives

$$M(\tfrac{1}{2}\pi, y) \leq K_3 y + K_4 .$$

Substituting in (12) we get the required result.

4. Theorems A and B are easily restated for the case where the region under consideration is the strip $x_1 < x < x_2$, by mapping this strip on G. In this section of our paper $G(x_1, x_2)$ is the region $x_1 < x < x_2$, $-\infty < y < \infty$. We shall assume, for convenience, that $0 < x_1 < 1 < x_2$. The boundary of $G(x_1, x_2)$ is $\Gamma(x_1, x_2)$ which is composed of $\Gamma_1(x_1, x_2)$: $x = x_1$ and of $\Gamma_2(x_1, x_2) : x = x_2$. The Green's function is $g(\zeta, z; x_1, x_2)$ and, if $f(z)$ is of bounded type in $G(x_1, x_2)$, $U_1(z; x_1, x_2)$ is the smallest positive harmonic function in $G(x_1, x_2)$ which $\geq \log | f(z) |$. When $x_2 = \infty$, we write $G(x_1)$, etc., and when, in addition, $x_1 = 0$, we write $\mathbf{G}$, etc. In adapting the theory of the preceding paragraphs for the region $G(x_1, x_2)$ we replace $g(\zeta, z)$ by $g(\zeta, z; x_1, x_2)$ throughout. If $k = \pi/(x_2 - x_1)$ and $\zeta' = k(\zeta - x_1)$, $z' = k(z - x_1)$, we have $g(\zeta', z') = g(\zeta, z; x_1, x_2)$. We find that

$$\eta_1 = \lim_{y \to \infty} e^{-k|y|} m(y), \qquad m(y) = \int_{x_1}^{x_2} \log^+ | f(x + iy) | \sin k(x - x_1) \, dx,$$

and that the inequality in (IV) becomes:

$$\text{(IV}'\text{)} \qquad \int_{-y}^{y} \log^+ |f(x+it)| e^{-k|t|}\, dt \le C_1 + C_2\, y \sin k(x - x_1) \qquad (x_1 < x < x_2).$$

THEOREM C. *Suppose that $f(z)$ is meromorphic in* $\mathbf{G}$. *Then $f(z)$ is of bounded type in* $\mathbf{G}$ *if and only if*

$$\text{(V)} \qquad \sum_{\mathbf{G}} \frac{\beta_\nu}{(\beta_\nu - 1)^2 + \gamma_\nu^2} < \infty \qquad (b_\nu \ne 1),$$

(VI) *a constant K exists such that*

$$\mathbf{m}(x) = \int_{-\infty}^{\infty} \log^+ |f(x+iy)| \frac{1}{(1+x)^2 + y^2}\, dy < K \qquad (x > 0).$$

If $f(z)$ is meromorphic in $G(x_1)$ and on $x = x_1$, and is of bounded type in $G(x_1)$, then the function $U_1(z; x_1)$ is given by

$$\text{(13)} \qquad U_1(x+iy; x_1) = \int_{-\infty}^{\infty} \log^+ |f(x_1+it)| \frac{x - x_1}{(x-x_1)^2 + (y-t)^2}\, dt + \xi(x - x_1) + \sum_{G(x_1)} g(x+iy, b_\nu; x_1),$$

where $\xi = \lim_{x\to\infty} \mathbf{m}(x)$.

The basis of our argument is that $f(z)$ is of bounded type in $\mathbf{G}$ if and only if (1) $f(z)$ is of bounded type in every $G(x_1, x_2) \subset \mathbf{G}$ and (2) $U_1(z; x_1, x_2)$ has a harmonic limit function when $G(x_1, x_2) \to \mathbf{G}$. It is actually sufficient that the various parts of $U_1(z; x_1, x_2)$ be bounded for $z = 1$, where the terms, if any, corresponding to $b_\nu = 1$ are dropped from the sum.

Using the same method that was used for condition I, we see that

$$\lim_{G(x_1,x_2)\to\mathbf{G}} \sum_{G(x_1,x_2)} g(z, b_\nu; x_1, x_2) = \sum_{\mathbf{G}} \mathbf{g}(z, b_\nu),$$

$$\mathbf{g}(\zeta, z) = \log\left|\frac{\zeta + z^*}{\zeta - z}\right| = \tfrac{1}{2}\log\frac{(s+x)^2 + (t-y)^2}{(s-x)^2 + (t-y)^2},$$

$$\mathbf{g}(1, b) = \tfrac{1}{2}\log\left[1 + \frac{4\beta}{(1-\beta)^2 + \gamma^2}\right],$$

which gives us condition V.

We asume that (V) and (VI) are fulfilled and show first that $f(z)$ is of bounded type in $G(x_1, x_2)$. Condition (I) is obviously satisfied, and since

$$e^{-k|y|} = e^{-\pi|y|/(x_2-x_1)} \le \frac{(x_2-x_1)^2}{(x_2-x_1)^2 + y^2},$$

(VI) implies that (IV′) holds in $G(x_1, x_2)$. Moreover, (IV′) must hold with $C_2 = 0$. Referring to (10), we see that this in turn implies that

$$\int_{-\infty}^{\infty} e^{-k|y|}\, m(y)\, dy$$

is finite, and hence $\eta_1 = \eta_2 = 0$.

Let $A_1(z; x_1, x_2)$ stand for the part of the integral in $U_1(z; x_1, x_2)$ which is taken over $\Gamma_1(x_1, x_2)$ and $A_2(z; x_1, x_2)$ for the rest of the integral. By comparing $g(\zeta, z; x_1, x_2)$ to $g(\zeta, z; x_1)$, one finds that

$$\frac{\partial}{\partial n} g(\zeta, 1; x_1, x_2)\,|_{\Gamma_1} \le \frac{1 - x_1}{(1 - x_1)^2 + t^2}.$$

Therefore, $A_1(1; x_1, x_2)$ is bounded for x_1 near 0.

We must show, finally, that $A_2(1; x_1, x_2)$ is bounded as $x_2 \to \infty$. By mapping $G(x_1, x_2)$ onto G we find that

$$\frac{\partial}{\partial n} g(\zeta, 1; x_1, x_2)\,|_{\Gamma_2(x_1, x_2)} = 2k\, \frac{\sin k(1 - x_1)}{e^{kt} + e^{-kt} + 2 \cos k(1 - x_1)}.$$

When $x_2 - x_1 > \pi$ $(k < 1)$, and x_1 is small, we get

$$\frac{\partial}{\partial n} g(\zeta, 1; x_1, x_2)\,|_{\Gamma_2(x_1, x_2)} \le 2k^2 e^{-k|t|} \le \frac{2\pi^2}{(x_2 - x_1)^2 + t^2}.$$

Condition (VI) tells us that $A_2(1; x_1, x_2)$ is bounded as $x_2 \to \infty$.

If $f(z)$ is of bounded type in $G(x_1)$ and is meromorphic on $x = x_1$, the representation (13) may be obtained by transforming $U_1(z)$, remembering that, since $f(z)$ is meromorphic on $x = x_1$, one of the constants η_i will be zero.

Assuming, now, that $f(z)$ is of bounded type in **G**, we wish to show that (VI) must be satisfied. Since $U_1(z; x_1)$ must exist for all x_1 and must be bounded for x_1 near zero, z fixed, it is easily seen that $\mathbf{m}(x)$ must be bounded for x_1 near zero. Suppose now that $x > x_1 > 0$. We have

$$\mathbf{m}(x) = \frac{\pi}{1 + x} \frac{1}{\pi} \int_{-\infty}^{\infty} \log^+ |f(x + it)|\, \frac{1 + 2x - x}{(1 + 2x - x)^2 + t^2}\, dt.$$

Using (13) with $z = 1 + 2x$ and with $x_1 = x$, we have

$$\frac{1 + x}{\pi}\, \mathbf{m}(x) = U_1(1 + 2x; x) - \xi(x)(1 + 2x - x) - \sum_{G(x)} g(1 + 2x, b_\nu; x).$$

We have written $\xi(x)$ in the above equation since we have not proved as yet that

ξ is independent of the region $G(x)$. Since $G(x) \subset G(x_1)$, $U_1(1 + 2x; x) \leq U_1(1 + 2x; x_1)$ and

$$\frac{1 + x}{\pi} \mathbf{m}(x) \leq \frac{1}{\pi} \int_{-\infty}^{\infty} \log^+ | f(x_1 + it) | \frac{1 + 2x - x_1}{(1 + 2x - x_1)^2 + t^2} dt$$

$$+ \xi(x_1)(1 + 2x - x_1) - \xi(x)(1 + x)$$

$$+ \sum_{G(x_1)} g(1 + 2x, b_\nu ; x_1) - \sum_{G(x)} g(1 + 2x, b_\nu ; x).$$

We wish to show that $\overline{\lim}_{x \to \infty} \mathbf{m}(x)$ is finite. We see that

$$\lim_{x \to \infty} \frac{1 + 2x - x_1}{1 + x} \int_{-\infty}^{\infty} \log^+ | f(x_1 + it) | \frac{1}{(1 + 2x - x_1)^2 + t^2} dt = 0,$$

$$\overline{\lim_{x \to \infty}} \frac{1}{1 + x} [\xi(x_1)(1 + 2x - x_1) - \xi(x)(1 + x)] = 2\xi(x_1) - \underline{\lim}_{x \to \infty} \xi(x)$$

$$\leq 2 \xi(x_1),$$

$$\sum_{G(x_1)} g(1 + 2x, b_\nu ; x_1) - \sum_{G(x)} g(1 + 2x, b_\nu ; x)$$

$$= \sum_{x_1 < \beta_\nu} \log \left| \frac{1 + 2x + b_\nu^* - 2x_1}{1 + 2x - b_\nu} \right| - \sum_{x < \beta_\nu} \log \left| \frac{1 + 2x + b_\nu^* - 2x}{1 + 2x - b_\nu} \right|$$

$$= \tfrac{1}{2} \sum_{x_1 < \beta_\nu \leq x} \log \frac{(1 + 2x + \beta_\nu - 2x_1)^2 + \gamma_\nu^2}{(1 + 2x - \beta_\nu)^2 + \gamma_\nu^2}$$

$$+ \tfrac{1}{2} \sum_{x < \beta_\nu} \log \frac{(1 + 2x + \beta_\nu - 2x_1)^2 + \gamma_\nu^2}{(1 + \beta_\nu)^2 + \gamma_\nu^2}$$

$$= \tfrac{1}{2} \sum_{x_1 < \beta_\nu \leq x} \log \left[1 + \frac{4(1 + 2x - x_1)(\beta_\nu - x_1)}{(1 + 2x - \beta_\nu)^2 + \gamma_\nu^2} \right]$$

$$+ \tfrac{1}{2} \sum_{x < \beta_\nu} \log \left[1 + \frac{4(1 + x + \beta_\nu - x_1)(x - x_1)}{(1 + \beta_\nu)^2 + \gamma_\nu^2} \right]$$

$$\leq 4(x + 1) \left[\sum_{x_1 < \beta_\nu \leq x} \frac{(\beta_\nu - x_1)}{(1 + 2x - \beta_\nu)^2 + \gamma_\nu^2} + \sum_{x < \beta_\nu} \frac{(1 + \beta_\nu)}{(1 + \beta_\nu)^2 + \gamma_\nu^2} \right].$$

The above series is convergent since $\sum \frac{\beta_\nu}{(\beta_\nu - 1)^2 + \gamma_\nu^2}$ is convergent. Therefore, $\mathbf{m}(x)$ must be bounded as x becomes infinite.

It is now easy to show that, if $f(z)$ is of bounded type, $\lim_{x\to\infty} \mathbf{m}(x) = \xi$. In the convergent series above, the ν-th term is constant when $x < \beta_\nu$ and then decreases to zero as x increases. Therefore, the whole series approaches zero and

$$\overline{\lim_{x\to\infty}} \frac{1}{\pi} \mathbf{m}(x) \leq 2\xi(x_1) - \underline{\lim_{x\to\infty}}\ \xi(x).$$

Letting x_1 become infinite, we get $\overline{\lim_{x\to\infty}} \frac{1}{\pi} \mathbf{m}(x) \leq \underline{\lim_{x\to\infty}}\ \xi(x)$. Substituting back, we get

$$\overline{\lim_{x\to\infty}} \frac{1}{\pi} \mathbf{m}(x) \leq \xi(x_1).$$

Suppose now that $\underline{\lim_{x\to\infty}} \frac{1}{\pi} \mathbf{m}(x) = \xi < \xi(x_1)$. The function $U_1(z; x_1) - \log | f(z) |$ is positive and harmonic in $G(x_1)$. Consider $m_1(x)$:

$$m_1(x) = \int_{-\infty}^{\infty} [U_1(x + iy) - \log | f(x + iy) |] \frac{1}{(1 + x)^2 + y^2} dy$$

$$\geq \int_{-\infty}^{\infty} \xi(x_1) \frac{(x - x_1)}{(1 + x)^2 + y^2} dy - \int_{-\infty}^{\infty} \log^+ | f(x + iy) | \frac{1}{(1 + x)^2 + y^2} dy$$

$$= \pi\xi(x_1) \frac{x - x_1}{1 + x} - \mathbf{m}(x),$$

so that

$$\overline{\lim_{x\to\infty}} \frac{1}{\pi} m_1(x) = \xi(x_1) - \xi.$$

That is, the "ξ" for $U_1 - \log | f(z) |$ is at least as big as $\xi(x_1) - \xi$, and we have:

$$U_1(x + iy; x_1) - \log | f(x + iy) | \geq [\xi(x_1) - \xi](x - x_1),$$

$$\log | f(x + iy) | \leq U_1(x + iy; x_1) - [\xi(x_1) - \xi](x - x_1).$$

The function on the right is positive, since $U_1(x + iy; x_1) \geq \xi(x_1)(x - x_1)$ and

is less than $U_1(z; x_1)$. But $U_1(z; x_1)$ is the smallest positive harmonic function in $G(x_1)$ which $\geq \log | f(z) |$. Therefore

$$\lim_{x \to \infty} \frac{1}{\pi} \mathbf{m}(x) = \xi(x_1) = \xi.$$

This completes the proof of Theorem C.

BIBLIOGRAPHY

1. L. V. AHLFORS, *On Phragmén-Lindelöf's principle*, Transactions of the American Mathematical Society, vol. 41(1937), pp. 1–8.
2. L. BIEBERBACH, *Lehrbuch der Funktionentheorie*, Leipzig and Berlin, 1931, vol. 2, p. 58.
3. E. HILLE, *On Laplace Integrals*, Attonde Skandinaviska Matematikerkongressen, Stockholm, 1934.
4. R. NEVANLINNA, *Ueber die Eigenschaften meromorpher Funktionen in einem Winkelraum*, Acta Societatis Scientiarum Fennicae, Stockholm, 1934.
5. R. NEVANLINNA, *Eindeutige Analytische Funktionen*, Berlin, 1936, chap. VII, §§1 and 2.

VASSAR COLLEGE.

DIFFERENTIALS AND ANALYTIC CONTINUATION IN NON-COMMUTATIVE ALGEBRAS

By R. W. Wagner

This paper represents an attempt to generalize the power series portion of the theory of functions of a complex variable so that the argument may range over a non-commutative algebra.

For this purpose, an absolute value, or metric, is defined for the algebra. This absolute value is used to discuss the convergence of generalized power series. By introducing differentials, a generalization of Taylor's theorem is obtained. This involves several theorems concerning the existence and differentiability of generalized power series. The final section of the paper consists of the application of this theory to the special case of functions defined on the complete matrix algebra.

One of the important differences between this generalization and the classical theory is that the region of convergence and the region of absolute convergence are very different when a general algebra is considered. This complicates the theory of convergence of series, but is a blessing in the analytic continuation of the function.

German letters will be used to denote sets of elements; capital letters denote elements of the algebra; and small letters will denote elements of the field used to build the algebra.

1. **The metric.** Let $\mathfrak{A}$ be a linear associative algebra over the field of all real numbers $\mathfrak{R}$, or the field of all complex numbers $\mathfrak{C}$. These fields are chosen because they are complete and the proof of the existence of limits is greatly simplified. Besides, each of these fields has a commonly used absolute value which can be extended to $\mathfrak{A}$.

An absolute value for $\mathfrak{A}$ should have the properties:

$$|X| \geq 0; \qquad |X| = 0 \text{ if, and only if, } X = 0; \tag{1.1}$$

$$|X + Y| \leq |X| + |Y|; \tag{1.2}$$

$$|XY| \leq |X| \, |Y|; \tag{1.3}$$

$$|xY| = |x| \, |Y|; \tag{1.4}$$

$$|I| = 1, \text{ if } \mathfrak{A} \text{ has an identity;} \tag{1.5}$$

$$|X| \text{ is continuous in coördinates for basis of } \mathfrak{A}. \tag{1.6}$$

Received March 8, 1941; in revised form June 30, 1942.

The absolute value to be used here is obtained by applying an absolute value function for matrices, which has been studied by Rella [6], to the regular representation of $\mathfrak{A}$. Let R_X be the matrix which represents X in the first matrix representation of $\mathfrak{A}$. Define $| R_X |$ as the positive square root of the largest characteristic root of the matrix $R_X R_X^*$, where R_X^* is the conjugate of R_X transposed. Now set

$$(1.7) \qquad | X | = | R_X |.$$

This gives an absolute value function for $\mathfrak{A}$. Rella has shown that this absolute value for matrices satisfies all the above conditions except (1.5). It is easy to verify that this absolute value satisfies this condition also.

Rella mentions other absolute value functions for matrices. But they fail to satisfy (1.1) or (1.5). The absolute value for an algebra used by Spampinato [9] fails to satisfy (1.3) and (1.5). When it is adjusted so that it will satisfy one of these, it may still fail to satisfy the other.

The absolute value used here is invariant under a change of basis by an orthogonal matrix, but not for a general change of basis. The other absolute values which were mentioned share this disadvantage. This absolute value is the norm of the linear transformation of a Euclidean vector space which the matrix represents.

The distance from A to B is $| A - B |$. A change of basis gives a metric which is topologically equivalent to the old one.

2. **Generalized power series.** Henceforth, the word "function" will be used to denote a correspondence between the elements of two subsets of $\mathfrak{A}$.

A multilinear function is a polynomial which is homogeneous of the first degree in each of k independent variables. For example, $AXBYC + BYAXC$ is a bilinear function of X and Y, with A, B, and C constants. If all the letters are variables, it is a multilinear function of degree 5. More generally,

$$(2.1) \qquad G_k(X_1, X_2, \cdots, X_k) = \sum A_0 X_1 A_1 X_2 A_2 \cdots X_k A_k$$

is a multilinear function of degree k. It is to be understood that the X_i may appear in different orders in different terms and that the multipliers A_i may vary from term to term.

Associated with the multilinear function (2.1) is a number g_k, called the norm of the multilinear function. This norm is defined to be the g.l.b. of numbers m such that

$$(2.2) \qquad | G_k(X_1, X_2, \cdots, X_k) | \leq m, \qquad \text{with } | X_i | = 1.$$

If all the variables have the same value, say $X_i = X$, the multilinear function becomes a homogeneous polynomial of k-th degree in X. Such a polynomial will be denoted by $G_k(X)$. From (2.2), one gets

$$(2.3) \qquad | G_k(X) | \leq g_k | X |^k.$$

It may well be that g_k is not the smallest number satisfying (2.3). But in some of the subsequent proofs it is necessary to have an inequality which may also be applied to a multilinear function. For uniformity, we use the same norm for both the homogeneous polynomial of degree k and the multilinear function.

A generalized power series is a sum of homogeneous polynomials in X. If $F(X)$ is a generalized power series,

$$F(X) = \sum_{k=0}^{\infty} G_k(X). \tag{2.4}$$

The theory of convergence of these series is complicated by the possible existence of divisors of zero in $\mathfrak{A}$. If X is a properly nilpotent element of $\mathfrak{A}$, all but a finite number of terms of (2.4) are zero and convergence is automatic. The non-commutativity of A also complicates the situation. Consider the two series

$$F_1(X) = \sum_{k=0}^{\infty} A_k X^k, \qquad F_2(X) = \sum_{k=0}^{\infty} X^k A_k,$$

where

$$A_k = \begin{pmatrix} 1 & 2^k \\ 0 & 1 \end{pmatrix}, \qquad X = \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix}.$$

These series may be written in the form

$$F_1(X) = \sum_{k=0}^{\infty} \begin{pmatrix} x^k & 2^k y^k \\ 0 & y^k \end{pmatrix}, \qquad F_2(X) = \sum_{k=0}^{\infty} \begin{pmatrix} x^k & 2^k x^k \\ 0 & y^k \end{pmatrix}.$$

Now it is clear that, if $x = \frac{3}{4}$ and $y = \frac{1}{4}$, the first series will converge but the second will diverge.

Some of these difficulties may be avoided by considering absolute convergence, or convergence of the series of absolute values.

THEOREM 2.1. *If a series converges absolutely, in the sense of the absolute value defined above, it converges.*

Let a_k denote the absolute value of the k-th term of the series in (2.4) and let S_r denote the sum of the first r terms of the series. Then

$$| S_{r+p} - S_r | \leq \sum_{k=r+1}^{r+p} a_k.$$

By hypothesis, for any positive e there exists an m such that $r \geq m$ implies

$$\sum_{k=r+1}^{r+p} a_k \leq e.$$

Therefore, $| S_{m+p} - S_m | \leq e$. Use E_i $(i = 1, 2, \cdots, n)$ for a basis of $\mathfrak{A}$ and write

$$S_r = \sum_{i=1}^{n} s_{r,i} E_i .$$

Then, for every p,

$$| \sum_{i=1}^{n} (s_{m+p,i} - s_{m,i})E_i | \leq e. \tag{2.5}$$

It will now be shown that there exists a number b, which depends only upon the basis, such that (2.5) implies

$$| s_{m+p,i} - s_{m,i} | \leq be \qquad (i = 1, 2, \cdots, n). \tag{2.6}$$

Let b_i be the minimum, with respect to the x_j, of

$$| (x_i + 1)E_i - \sum_{j=1}^{n} x_j E_j |.$$

No $b_i = 0$. For, if it did, one would have, by (1.1), a linear dependence of the E_j, which is impossible. Let b be the reciprocal of the smallest b_i. Then, for every i and arbitrary values of x_j, one gets

$$| \sum_{j=1}^{n} x_j E_j | = | x_i | \ | \sum_{j=1}^{n} x_j E_j / x_i | \geq | x_i | / b.$$

Therefore, $| x_i | \leq b | \sum x_j E_j |$. Applying this result to (2.5), one gets (2.6). Because the coefficient field is complete, (2.6) implies that for every i there exists a number s_i which is the limit of the sequence $s_{m,i}$. Therefore, the sequence S_r converges to $\sum s_i E_i$.

The region of absolute convergence of the series will usually be much smaller than the region of convergence. In the situation considered here, the closure of the region of absolute convergence is, in general, a proper subset of the closure of the region of convergence. In the case of a complex variable, the closures of these regions are the same.

It is usually very difficult to find, or define in terms of the coördinates, a maximal region of absolute convergence. It may also be difficult to define a maximal region of absolute convergence in terms of absolute values. However, it is sometimes very simple to find a region in which the convergence is absolute. The following theorem gives a very crude, but sometimes useful, criterion for the absolute convergence of a generalized power series.

THEOREM 2.2. *If* $\sum g_k x^k$ *converges on* $| x | < c$, *the series* (2.4) *will converge absolutely and uniformly on* $| X | \leq c' < c$.

By applying (1.2) and (2.2), one gets

$$| \sum_{k=0}^{\infty} G_k(X) | \leq \sum_{k=0}^{\infty} | G_k(X) | \leq \sum_{k=0}^{\infty} g_k | X |^k.$$

Therefore the series converges absolutely on $| X | < c$. The Heine-Borel theorem may be applied to the closed region $| X | \leq c'$, so it is easy to prove uniformity of convergence.

3. **Differentials.** Let $F(X)$ be a function which may or may not have a generalized power series. The differential [1] of this function, if it exists, is a function of X and of the increment of X (call this increment H) which is linear in the increment and, in a sense, approximates the function. The differential (denote it by $D[F(X), H]$) is the linear and homogeneous function of H with the property that for any $e > 0$ there exists a d such that $|H| \leq d$ implies

$$|F(X + H) - F(X) - D[F(X), H]| \leq e|H|.$$

From this relation it is easy to prove the following theorems.

THEOREM 3.1. *If the function has a differential, this differential is unique.*

THEOREM 3.2. *If each of two functions has a differential, the differential of the sum is the sum of the differentials.*

THEOREM 3.3. *If each of two functions has a differential, the differential of their product is the differential of the first multiplied by the second function plus the first function multiplied by the differential of the second.*

This differential has been used extensively in general analysis. The existence of this differential is more restrictive than the existence of the Hamiltonian differential as defined in Wedderburn [12]. The Hamiltonian differential is directional. In order for the Fréchet differential (defined above) to exist, all directional differentials must exist and it must be possible to incorporate them all into a single linear function defined over the whole space. Graves (see [2] or [3]) clarifies this distinction in a more general case.

Higher differentials have been defined by Graves (see [2] or [3]). The k-th differential of a function is a multilinear function of k independent increments. It is defined inductively. It satisfies this condition. If $e > 0$, there exists a d such that $|H_i| \leq d$ implies that

$$|F(X + H_1 + \cdots + H_k) - F(X + H_1 + \cdots + H_{k-1}) - D^k[F(X), H_1 \cdots H_k]| \leq e|H_1| \cdots |H_k|.$$

For our purposes it is convenient to specialize this by making all the increments equal after finding this multilinear function. Then the k-th differential is denoted by $D^k[F(X), H]$.

THEOREM 3.4. *If the r-th differential exists, it is unique.*

THEOREM 3.5. *The r-th differential of a sum of functions, each of which has an r-th differential, is the sum of the r-th differentials.*

The proofs of these theorems are omitted. Some special cases are of interest, for they show how these differentials compare with the derivatives which are used in commutative cases. For example,

$$D[X^3, H] = X^2H + XHX + HX^2,$$

and, if $G_k(X)$ is a homogeneous polynomial in X of degree k,

$$D^k[G_k(X), H] = k!\, G_k(H). \tag{3.1}$$

THEOREM 3.6. *If $F(X)$ is defined by (2.4) and if $\sum g_k \mid X \mid^k$ converges uniformly on $\mid X \mid \leq c$, then $\sum D[G_k(X), H]$ converges absolutely and uniformly on $\mid X \mid \leq c$ and equals $D[F(X), H]$.*

First, it will be shown that the sum of the differentials converges uniformly. By applying Theorem 3.3 and (2.2) one gets

$$\mid D[G_k(X), H] \mid \leq kg_k \mid H \mid \; \mid X \mid^{k-1}.$$

The hypothesis that $\sum g_k \mid X \mid^k$ converges uniformly on $\mid X \mid \leq c$ implies that $\sum kg_k \mid H \mid \; \mid X \mid^{k-1}$ converges uniformly on $\mid X \mid \leq c$. Therefore, the sum of the differentials converges uniformly.

The proof that the sum of the differentials is $D[F(X), H]$ is separated into two cases. First, suppose $X = 0$. Then one gets

$$F(0 + H) - F(0) - G_1(H) = \sum_{k=2}^{\infty} G_k(H),$$

and taking absolute values, one gets, by using (2.2),

$$\mid F(0 + H) - F(0) - G_1(H) \mid \leq \mid H \mid^2 [\sum_{k=2}^{\infty} g_k \mid H \mid^{k-2}].$$

The quantity in brackets has a bound on the region $\mid H \mid \leq c$. Call this bound b. Take $\mid H \mid$ less than the smaller of c and e/b. Then surely

$$\mid F(0 + H) - F(0) - G_1(H) \mid \leq e \mid H \mid.$$

Therefore, $G_1(H) = D[F(0), H]$. Moreover, if $k > 1$, $D[G_k(0), H] = 0$. Therefore, $\sum D[G_k(0), H] = D[F(0), H]$.

Now consider the case with $X \neq 0$. First, one has

$$\mid G_k(X + H) - G_k(X) \mid \leq g_k \mid H \mid [\sum_{s=1}^{k} X^{k-s} H^{s-1}].$$

Now one can assume that $\mid H \mid \leq \mid X \mid$, so that this reduces to

$$\mid G_k(X + H) - G_k(X) \mid \leq kg_k \mid H \mid \; \mid X \mid^{k-1}.$$

For any $e > 0$ there exists an m such that

$$\mid \sum_{k=m}^{\infty} G_k(X + H) - G_k(X) \mid \leq \mid H \mid \sum_{k=m}^{\infty} kg_k \mid X \mid^{k-1} \leq e \mid H \mid \tag{3.2}$$

and such that

$$\mid \sum_{k=m}^{\infty} D[G_k(X), H] \mid \leq \mid H \mid \sum_{k=m}^{\infty} kg_k \mid X \mid^{k-1} \leq e \mid H \mid, \tag{3.3}$$

provided $\mid X \mid \leq c$.

One also has in this case

$$| F(X + H) - F(X) - \sum_{k=0}^{m-1} [G_k(X + H) - G_k(X)] |$$
$$\leq \sum_{k=m}^{\infty} | G_k(X + H) - G_k(X) |,$$

provided $| H | \leq | X |$ and $| H | \leq c - | X |$. Using (3.2), one gets

$$(3.4) \qquad | F(X + H) - F(X) - \sum_{k=0}^{m-1} [G_k(X + H) - G_k(X)] | \leq e | H |.$$

However, a finite sum has a differential; therefore there exists a d such that $| H | \leq d$ implies

$$(3.5) \qquad | \sum_{k=0}^{m-1} [G_k(X + H) - G_k(X)] - \sum_{k=0}^{m-1} D[G_k(X), H] | \leq e | H |.$$

When one combines (3.3), (3.4), and (3.5), he gets

$$| F(X + H) - F(X) - \sum_{k=0}^{\infty} D[G_k(X), H] | < 3e | H |$$

and this is valid provided $| H |$ is less than the smallest of the quantities $| X |$, $c - | X |$, and d. Therefore, the sum of the differentials of the polynomials is $D[F(X), H]$.

The preceding theorem may be stated and proved in the same fashion by using r-th differentials.

Let $E_1, E_2, \cdots, E_n$ denote a set of basis elements for $\mathfrak{A}$. Then, if $Y = F(X)$, one can write

$$X = \sum_{r=1}^{n} x_r E_r, \qquad H = \sum_{r=1}^{n} h_r E_r, \qquad Y = \sum_{s=1}^{n} y_s E_s,$$

where y_s is a function of the independent variables $x_1, x_2, \cdots, x_n$ in the usual sense.

THEOREM 3.7. *If $D^k[Y, H]$ exists, then*

$$D^k[Y, H] = \sum_{s=1}^{n} \sum \frac{\partial^k y_s E_s}{\partial x_{r_1} \partial x_{r_2} \cdots \partial x_{r_k}} h_{r_1} h_{r_2} \cdots h_{r_k},$$

where the summation extends to all the n^k ways of choosing the subscripts r_i.

The proof is given only in the case $k = 1$. Apply Theorems 3.2 and 3.3 to the assumed differential to get

$$| E_s y_s(X + H) - E_s y_s(X) - E_s D[y_s(X), H] | \leq e | H |,$$

provided $| H |$ is sufficiently small. With $H = h_r E_r$ this may be written as

$$\left| E_s \left\{ \frac{y_s(X + E_r h_r) - y_s(X)}{h_r} - D[y_s(X), E_r] \right\} \right| \leq e \mid E_r \mid.$$

Therefore, $D[y_s(X), E_r]$ is the partial derivative of y_s with respect to x_r. Recall that $D[Y, H]$ is linear and homogeneous in H. Therefore,

$$D[y_s(X), H] = \sum_{r=1}^{n} \frac{\partial y_s}{\partial x_r} h_r .$$

By applying Theorems 3.2 and 3.3 again, one gets the result stated for this case. For other values of k, the proof follows the same lines.

4. **Taylor's theorem and generalized power series.** In this section the relationships between differentials and generalized power series are discussed. The first step is to consider homogeneous polynomials.

THEOREM 4.1. $G_k(X + H) = \sum_{r=0}^{k} D^r[G_k(X), H]/r!$.

First, let $G_k(X) = A_0 X A_1 X A_2 \cdots X A_k$. The theorem will be proved by induction for this particular function and then extended to the general homogeneous polynomial. Start with

$$G_1(X + H) = G_1(X) + G_1(H) = G_1(X) + D[G_1(X), H].$$

Therefore, the theorem is true if $k = 1$. Assume that it is true if $k = s - 1$. Then $G_s(X) = G_{s-1}(X) X A_s$. Using Theorem 3.3 several times, one gets

$$D^r[G_s(X), H] = D^r[G_{s-1}(X), H] X A_s + r D^{r-1}[G_{s-1}(X), H] H A_s .$$

Divide through by $r!$ and add the resulting equations for various values of r. The result may be written in the form

$$\sum_{r=0}^{s} D^r[G_s(X), H]/r! = G_{s-1}(X + H)[X + H] A_s = G_s(X + H).$$

Therefore, the theorem is true for this particular function. Any homogeneous polynomial of degree k is a sum of functions of this kind. Therefore, by Theorem 3.5, the result of this theorem is valid for any homogeneous polynomial.

THEOREM 4.2. *If $F(X)$ has a generalized power series expansion in a neighborhood of $X = A$, then*

$$F(A + H) = \sum_{r=0}^{\infty} D^r[F(A), H]/r!.$$

Suppose that the generalized power series is $\sum G_k(H)$. By the extension of Theorem 3.6, which was mentioned but not proved, $F(X)$ has differentials of all orders at $X = A$. Take the r-th differential of both members of the identity

$$F(A + H) = \sum_{k=0}^{\infty} G_k(H)$$

with H as the independent variable. The result is

$$D^r[F(A + H), K] = \sum_{k=0}^{\infty} D^r[G_k(H), K].$$

Now set $H = 0$. If $k \neq r$, $D^r[G_k(0), K] = 0$. Therefore,

$$D^r[F(A), K] = D^r[G_r(0), K] = r!\, G_r(K).$$

The last step depends upon (3.1). Thus the terms in the generalized power series are identified and the theorem is proved.

THEOREM 4.3. *Necessary and sufficient conditions that $F(X)$ have a generalized power series are that each y_s have a power series in the x_r, and that, for all values of r, $D^r[F(X), H]$ can be written as a homogeneous polynomial in H—the multipliers will depend upon X.*

Necessity. Since the generalized power series converges, it will converge in each of the directions represented by the basis elements. Therefore, each y_s will have a power series expansion in the x_r, and this series will represent this component of the function. Theorem 4.2 states the necessity of the second condition.

Sufficiency. Choose c' so that on the region $\sum x_r x_r^* \leq c'^2$, where x^* denotes the conjugate of x, each power series for y_s in the x_r converges absolutely. This region contains another, namely $| H | \leq c$, so that

$$y_s = \sum_{r=0}^{\infty} \sum \frac{\partial^r y_s}{r!\partial x_{k_1} \partial x_{k_2} \cdots \partial x_{k_r}} h_{k_1} h_{k_2} \cdots h_{k_r}$$

on the region $| H | \leq c$. The summation is to be extended to all the n^r ways of choosing the subscripts k_i. These series converge absolutely, so that, after multiplying both sides by E_s and adding for all values of s, one gets

$$Y = \sum_{r=0}^{\infty} \frac{1}{r!} \sum \sum_{s=1}^{n} \frac{E_s \partial^r y_s}{\partial x_{k_1} \partial x_{k_2} \cdots \partial x_{k_r}} h_{k_1} h_{k_2} \cdots h_{k_r}.$$

By Theorem 3.7, this may be written as

$$Y = \sum_{r=0}^{\infty} D^r[F(X), H]/r!$$

and the right member converges on the region $H \leq c$. By hypothesis, each differential may be replaced by a homogeneous polynomial of degree r in H. Thus the theorem is established.

This last condition is by no means trivial. It may be restated in this form: n polynomials in the variables x_r cannot always be written as a single polynomial in X. Whether they can or not depends upon the properties of the algebra $\mathfrak{A}$. For example, in the case of functions of an ordinary complex variable, $x^2 + ixy$ cannot be written as a polynomial in $z = x + iy$. In other words, this condition requires a kind of "analyticity" of the function.

Papers by Ward [11] and Ringleb [7] indicate that the answer to this question may be found by generalizing the Cauchy-Riemann differential equations. It appears to the author that such a generalization must include partial differential equations of order higher than the first in order to be of value in this connection.

THEOREM 4.4. *Let $F(X)$ be defined by a generalized power series which converges uniformly and absolutely on $|X| \leq 2c$. Then, on the region $|X| < c$, $|H| < c$,*

$$F(X + H) = \sum_{r=0}^{\infty} D^r[F(X), H]/r!.$$

From Theorem 4.1, one gets

$$F(X + H) = \sum_{k=0}^{\infty} G_k(X + H) = \sum_{k=0}^{\infty} \sum_{r=0}^{k} D^r[G_k(X), H]/r!. \tag{4.1}$$

This double series converges absolutely. For,

$$|D^r[G_k(X), H]/r!| \leq {}_kC_r g_k |X|^{k-r} |H|^r \leq {}_kC_r g_k c^k,$$

and from this relation one gets

$$\left| \sum_{k=0}^{\infty} \sum_{r=0}^{k} D^r[G_k(X), H]/r! \right| \leq \sum_{k=0}^{\infty} \sum_{r=0}^{k} {}_kC_r g_k c^k = \sum_{k=0}^{\infty} g_k(2c)^k.$$

This last series is convergent by hypothesis. Therefore, the terms of the double series may be rearranged. When one changes the order of summing in (4.1), he gets

$$F(X + H) = \sum_{r=0}^{\infty} \sum_{k=0}^{r} D^r[G_k(X), H]/r! = \sum_{r=0}^{\infty} D^r[F(X), H]/r!.$$

This last theorem enables one to shift the center of a "sphere" of absolute convergence. The new series, formed by such a shift, was proved to converge only on the region $|H| < c$. However, it may happen that the new series will converge on a region which extends outside the region of convergence of the original series. In this case one would have an analytic extension of the original series. An example of this is given in the next section.

5. **Matrix functions.** A vast number of papers has been published concerning the matrix functions which one gets by changing the variable in an ordinary power series to a matrix. (For a complete bibliography of this subject, see, for

instance, [11], [12], or [5].) The foregoing theory can be applied to these functions to get some interesting results.

Let $\mathfrak{M}$ be the complete matrix algebra of order n^2 over the field $\mathfrak{C}$. One way to define a matrix function is to change the power series

$$(x) = \sum_{k=0}^{\infty} a_k x^k \tag{5.1}$$

into the matrix function

$$F(X) = \sum_{k=0}^{\infty} a_k X^k. \tag{5.2}$$

Another method for generating a matrix function is to start with a set of partial idempotent elements, P_i, and the associated characteristic roots, λ_i, and nilpotent elements, Q_i $(i = 1, 2, \cdots, \nu)$. Then, if $g(x)$ is an analytic function, and not just an element of such a function, one can write a matrix function in Schwerdtfeger's form [8]

$$G(X) = \sum_{i=1}^{\nu} \sum_{r=0}^{n} g^{(r)}(\lambda_i)\, P_i Q_i^r \,/r!. \tag{5.3}$$

If $g(x)$ agrees with $f(x)$ on the region where $f(x)$ is defined, then $G(X)$ is an analytic continuation of $F(X)$ [10].

THEOREM 5.1. *If $G(X)$ is defined by* (5.3), *and if X is a regular point of $G(X)$, then*

$$D[G(X), H] = \sum_{i,j=1}^{\nu} \sum_{r,s=1}^{n} \frac{\partial^{r+s} g'(\lambda_i, \lambda_j)}{\partial\lambda_i^r\, \partial\lambda_j^s} P_i Q_i^r\, H P_j\, Q_j^s/r!s!, \tag{5.4}$$

where

$$g'(x, y) = [g(x) - g(y)]/(x - y). \tag{5.5}$$

The proof of this theorem consists in showing that the identity of the functions $F(X)$ and $G(X)$ can be extended to their differentials. The first step is to use Theorem 3.6 and to show that the differential of a matrix function defined by (5.2) is of the form shown in (5.4). This calculation is tedious and will be omitted. Since $G(X)$ and $F(X)$ are identical when both are defined, the same is true of their differentials. Therefore, on the region for which $F(X)$ is defined, the theorem is valid. The function appears in (5.3) and (5.4) only with a complex number for its argument. Therefore, the theorem will also be true on any region to which $g(x)$ may be extended by analytic continuation, i.e., the range on which $G(X)$ is defined.

The differential of a function shows the local properties of the mapping induced by the function. Thus the preceding theorem enables one to describe the local properties of the mappings of these matrix functions.

Because the matrices P_i are (algebraically) orthogonal idempotents in $\mathfrak{M}$, the subspaces

$$\mathfrak{S}_{rs} = P_r \mathfrak{M} P_s$$

are a set of linear subspaces whose linear extension is $\mathfrak{M}$.

THEOREM 5.2. *The invariant subspaces of the linear transformation defined by (5.4) are $\mathfrak{S}_{rs}$ and the associated scalar factor is $g'(\lambda_r, \lambda_s)$.*

From (5.4) and the multiplicative properties of the partial idempotent elements, one gets

$$D[G(X), \mathfrak{S}_{rs}] = \sum_{\rho,\sigma=0}^{n} \frac{\partial^{\rho+\sigma} g'(\lambda_r, \lambda_s)}{\rho!\sigma!\partial\lambda_r^\rho \partial\lambda_s^\sigma} P_r Q_r^\rho \mathfrak{S}_{rs} P_s Q_s^\sigma .$$

The right member is obviously contained in $\mathfrak{S}_{rs}$ and the one term with $\rho = \sigma = 0$ is the whole of $\mathfrak{S}_{rs}$. Therefore, each $\mathfrak{S}_{rs}$ is an invariant subspace of the differential. In case the Q_i are all zero (all elementary divisors of X are linear), the theorem is proved. In the contrary case, suppose that $Q_r^\alpha \neq 0$, $Q_s^\beta \neq 0$ and that $Q_r^{\alpha+1} = Q_s^{\beta+1} = 0$. Then the subspace $Q_r^\alpha \mathfrak{S}_{rs} Q_s^\beta$ is invariant. For, one gets

$$D[G(X), Q_r^\alpha \mathfrak{S}_{rs} Q_s^\beta] = g'(\lambda_r, \lambda_s) P_r Q_r^\alpha \mathfrak{S}_{rs} P_s Q_s^\beta .$$

The associated scalar factor is $g'(\lambda_r, \lambda_s)$ and the theorem is proved.

THEOREM 5.3. *If all the nilpotent elements of X are zero and if X is a regular point of $G(X)$, the k-th differential of $G(X)$ is*

$$k! \sum_{i_0=1}^{r} \cdots \sum_{i_k=1}^{r} g^{(k)}(\lambda_{i_0}, \lambda_{i_1}, \cdots, \lambda_{i_k}) P_{i_0} H P_{i_1} H \cdots H P_{i_k} , \tag{5.6}$$

where $g^{(k)}(x_0, x_1, \cdots, x_k)$ is the k-th divided difference of $g(x)$.

The proof of this theorem follows the same pattern as that of Theorem 5.1 and is omitted. The restriction that the Q_i are zero is not essential to the proof but is inserted to simplify the resulting expression for the differential.

THEOREM 5.4. *If all nilpotent elements of X are zero and if all the characteristic roots lie within a circle of convergence of an analytic element of $g(x)$, there is a neighborhood of X on which*

$$G(X + H) = \sum_{r=0}^{\infty} D^r[G(X), H]/r!.$$

First, it will be shown that there is a neighborhood of X on which the sum of the differentials converges absolutely. Let $\sum a_k(x - a)^k$ be the analytic element of $g(x)$ and let $g^*(x)$ be the function $\sum | a_k | \; | x - a |^k$. Let c be the radius of convergence of these two series. Then, from

$$g^{(r)}(\lambda_{i_0}, \cdots, \lambda_{i_r}) = \sum_{s=0}^{\infty} \sum{}' a_{r+s}(\lambda_{i_0} - a)^{\alpha_0} \cdots (\lambda_{i_r} - a)^{\alpha_r},$$

where $\sum'$ means to add all such terms with $\sum \alpha_i = s$, one gets

$$| g^{(r)}(\lambda_{i_0}, \cdots, \lambda_{i_r}) | \leq \sum_{s=0}^{\infty} \sum{}' \, | a_{r+s} | \; | \lambda_{i_0} - a |^{\alpha_0} \cdots | \lambda_{i_r} - a |^{\alpha_r}.$$

Let m be the largest of the quantities $| \lambda_{i_0} - a |$. There are ${}_{r+s}C_s$ terms in $\sum'$. Therefore, one gets

$$| g^{(r)}(\lambda_{i_0}, \cdots, \lambda_{i_r}) | \leq \sum_{s=0}^{\infty} | a_{r+s} | \, {}_{r+s}C_s m^s.$$

However, from the definition of $g^*(x)$, one has

$$(r + s)! \, | a_{r+s} | = g^{*(r+s)}(a).$$

Finally one gets, then, the inequality

$$| g^{(r)}(\lambda_{i_0}, \cdots, \lambda_{i_r}) | \leq g^{*(r)}(m + a)/r!. \tag{5.7}$$

From Theorem 5.3 one has

$$| D^r[G(X), H]/r! | \leq \sum_{i_0=1}^{\nu} \cdots \sum_{i_k=1}^{\nu} g^{(r)}(\lambda_{i_0}, \cdots, \lambda_{i_r}) \, p^{r+1} \, | H |^r,$$

where p is the largest of the quantities $| P_i |$. Apply (5.7) to the right member of this inequality and add for various values of r. The result is

$$\sum_{r=0}^{\infty} | D^r[G(X), H]/r! | \leq \sum_{r=0}^{\infty} g^{*(r)}(m + a)(pn)^{r+1} \, | H |^r/r!. \tag{5.8}$$

The point $x = m + a$ is a regular point of $g^*(x)$ because $m < c$. Therefore, $g^*(x)$ has a power series expansion in $(x - m - a)$. Let c' denote the radius of convergence of this series. Then if $np \, | H | \leq b$, where b is less than the smaller of c' and $c - m$, the series in (5.8) converge absolutely.

In order to show that the sum of the differentials converges to $G(X + H)$, consider the related functions $f(x)$ and $F(X)$ in (5.1) and (5.2). Theorem 4.4 states the desired result for the function $F(X)$. Since $F(X + H)$ can also be written in the form (5.3), one gets

$$F(X + H) = \sum_{r=0}^{\infty} D^r[F(X), H]/r! = \sum_{i=1}^{\nu'} \sum_{r=0}^{n} f^{(r)}(\lambda_i')P_i'Q_i'^r/r!,$$

where λ_i', ν', P_i', and Q_i' refer to $X + H$. But Theorems 5.1 and 5.3 show how to write the differentials so that the function appears only as $f(x)$. Therefore, this identity will be valid for any analytic continuation of $f(x)$. Thus it is true for $g(x)$ and consequently for $G(X)$. This completes the proof of the theorem.

Example. Let $G(X) = (I + X)^{-1}$. A series for this function which is valid in a neighborhood of the origin is

$$G(X) = I - X + X^2 - X^3 + \cdots. \tag{5.9}$$

A generalized power series for $G(X)$, which is valid in a neighborhood of $X = B$, is

$$G(B + H) = A - AHA + AHAHA - AHAHAHA + \cdots, \tag{5.10}$$

where $A = (I + B)^{-1}$. Observe that this uses the formula

$$D[(I + B)^{-1}, H] = -(I + B)^{-1}H(I + B)^{-1}.$$

Let B represent the matrix

$$B = \frac{1}{8}\begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix}$$

and find $| B | = .41$. Therefore, (5.9) converges absolutely when $X = B$. With $H = I$, and this same value for B, (5.9) will diverge for $X = B + H = B + I$. With this value of H, the series (5.10) will converge absolutely for

$$| A | = | (I + B)^{-1} | = .82 < 1.$$

Therefore, (5.10) gives a value for $G(B + I)$ while (5.9) does not. Thus (5.10) represents an analytic extension of (5.9). This illustrates the situation described at the end of the preceding section.

Relation to the literature. Hausdorff [4] made one of the first attempts to get a generalization similar to the one given here. He used only the first differential and called a function analytic if its first differential existed and could be written in terms of the algebra. In this notation,

$$D[F(X), H] = \sum_{i,j=1}^{n} a_{ij}\, E_i H E_j\, .$$

Ringleb [7] extended this idea by showing how certain decompositions of the algebra are reflected in the function. He also tried to use expressions like

$$\sum a_{i_1, i_2, \cdots, i_k} x_1^{i_1} x_1^{i_2} \cdots x_k^{i_k}$$

to get generalized power series. He was unsuccessful because he used only the first differential. Ward [11] defined a derivative, even in the non-commutative case, but he did not get a satisfactory second derivative. Most other papers do not go very far until they make a restriction that the algebra is commutative or that any multipliers appear only on one side of the variable.

Bibliography

1. M. Fréchet, *La notion de différentielle dans l'analyse générale*, Annales Scientifiques de l'Ecole Normale Supérieure, (3), vol. 42(1925), pp. 293–323.
2. L. M. Graves, *Riemann integration and Taylor's theorem in general analysis*, Transactions of the American Mathematical Society, vol. 29(1927), pp. 163–177.
3. L. M. Graves, *Topics in the functional calculus*, Bulletin of the American Mathematical Society, vol. 41(1935), pp. 641–662.

4. F. HAUSDORFF, *Zur Theorie der Systeme complexer Zahlen*, Berichte über die Verhandlungen der Sächsischen Akademie der Wissenschaften zu Leipzig, Mathematisch-physikalische Klasse, vol. 52(1900), pp. 43–61.
5. C. C. MACDUFFEE, *The Theory of Matrices*, Berlin, 1933.
6. T. RELLA, *Über positiv-homogene Funktionen ersten Grades einer Matrix*, Monatshefte für Mathematik und Physik, vol. 48(1939), pp. 84–95.
7. F. RINGLEB, *Beiträge zur Funktionentheorie in Hypercomplexen Systemen* I, Rendiconti del Circolo Matematico di Palermo, vol. 57(1933), pp. 311–340.
8. H. SCHWERDTFEGER, *Ueber mehrdeutige Matrixfunktionen*, Compositio Mathematica, vol. 3(1936), pp. 380–390.
9. N. SPAMPINATO, *Sulle funzioni di una variabile, in un'algebra complessa, ad n unità, dotata di modulo*, Rendiconti del Circolo Matematico di Palermo, vol. 57(1933), pp. 235–271.
10. R. W. WAGNER, *An extension of analytic functions to matrices*, American Journal of Mathematics, vol. 62(1940), pp. 380–390.
11. J. A. WARD, *A theory of analytic functions in linear associative algebras*, this Journal, vol. 7(1940), pp. 233–248.
12. J. H. M. WEDDERBURN, *Lectures on Matrices*, American Mathematical Society Colloquium Publications, 1934, p. 131.

OBERLIN COLLEGE.

SOME THEOREMS ON DOUBLE TRIGONOMETRIC SERIES

By George E. Reves and Otto Szász

1. **Introduction.** The present paper generalizes to two variables the following theorems:

(a) The Cantor-Lebesgue theorem [7; 267]. If $a_n \cos nx + b_n \sin nx \to 0$ in a set of positive measure, then $a_n \to 0$, $b_n \to 0$.

(b) The Fatou-Denjoy-Lusin theorem [1], [7; 131]. If

$$\sum | a_n \cos nx + b_n \sin nx |$$

converges in a set of positive measure, then $\sum (| a_n | + | b_n |) < \infty$.

(c) Two theorems of Szász [5; 376-378] on absolute convergence of Fourier series.

In some special cases and in a slightly different way such results are given in the Ph.D. thesis of Reves (Cincinnati, 1941).

2. **Generalization of the Cantor-Lebesgue theorem to two variables.**

Theorem I. *Let*

$$A_{mn}(x, y) = a_{mn} \cos mx \cos ny + b_{mn} \sin mx \cos ny + c_{mn} \cos mx \sin ny + d_{mn} \sin mx \sin ny. \tag{2.1}$$

If

$$\lim_{\substack{m\to\infty \\ n\to\infty}} A_{mn}(x, y) = 0 \tag{2.2}$$

in a two-dimensional point set E of measure $| E | > 0$, then

(2.3) $\rho_{mn} \equiv (a_{mn}^2 + b_{mn}^2 + c_{mn}^2 + d_{mn}^2)^{\frac{1}{2}} \to 0$ *as m and n tend to* ∞.

Any double limit is meant in the Pringsheim sense, i.e., m, n tend to ∞ simultaneously but independently.

Proof. We may dismiss the A_{mn} for which $\rho_{mn} = 0$. Let

$$A_{mn}(x, y) = \rho_{mn} B_{mn}(x, y),$$

so that

$$B_{mn}(x, y) = \frac{A_{mn}(x, y)}{(a_{mn}^2 + b_{mn}^2 + c_{mn}^2 + d_{mn}^2)^{\frac{1}{2}}}. \tag{2.4}$$

Received November 12, 1941; presented to the American Mathematical Society, December 29, 1941.

We first prove that

$$(2.5) \qquad | B_{mn}(x, y) | \le 1 \qquad \text{for all } (x, y).$$

We define p_{mn}, q_{mn}, $\phi = \phi_{mn}$, $\theta = \theta_{mn}$ by

$$(2.6) \qquad \begin{aligned} a_{mn} &= p_{mn} \cos \phi, & b_{mn} &= p_{mn} \sin \phi, & (p_{mn} \ge 0,\ q_{mn} \ge 0),\\ c_{mn} &= q_{mn} \cos \theta, & d_{mn} &= q_{mn} \sin \theta, & (0 \le \phi, \theta < 2\pi),\end{aligned}$$

so that

$$(2.7) \qquad p_{mn}^2 = a_{mn}^2 + b_{mn}^2, \qquad q_{mn}^2 = c_{mn}^2 + d_{mn}^2, \qquad \rho_{mn}^2 = p_{mn}^2 + q_{mn}^2,$$

and

$$(2.8) \qquad A_{mn}(x, y) = p_{mn} \cos (mx - \phi) \cos ny + q_{mn} \cos (mx - \theta) \sin ny.$$

From (2.8) we have

$$| A_{mn}(x, y) | \le p_{mn} | \cos ny | + q_{mn} | \sin ny |.$$

An application of Schwarz's inequality gives

$$| A_{mn}(x, y) | \le (p_{mn}^2 + q_{mn}^2)^{\frac{1}{2}} = \rho_{mn}.$$

This relation proves (2.5). By assumption (2.2),

$$\rho_{mn} B_{mn} \to 0 \qquad \text{for } (x, y) \text{ in } E.$$

Hence, by using a theorem of Egoroff [4; II, 144], corresponding to any $\epsilon > 0$ there exists an E_1 contained in E with $| E_1 | > | E | - \epsilon > 0$, such that

$$\rho_{mn} B_{mn} \to 0 \qquad \text{uniformly for all } (x, y) \text{ in } E_1.$$

From (2.5) a fortiori

$$\rho_{mn} B_{mn}^2 \to 0 \qquad \text{uniformly for all } (x, y) \text{ in } E_1.$$

Hence, we can integrate over E_1 and get

$$\rho_{mn} \iint_{E_1} B_{mn}^2(x, y)\, dx\, dy \to 0.$$

To prove that ρ_{mn} tends to zero as m and n tend to ∞, it is sufficient to prove that

$$I_{mn} \equiv \iint_{E_1} B_{mn}^2(x, y)\, dx\, dy$$

is bounded away from zero for all m, n sufficiently large, or what is the same thing

$$(2.9) \qquad \varliminf_{\substack{m \to \infty \\ n \to \infty}} I_{mn} > 0.$$

We have from (2.4), (2.7) and (2.8) that

$$I_{mn} = \iint_{E_1} \frac{1}{p_{mn}^2 + q_{mn}^2} [p_{mn}^2 \cos^2(mx - \phi) \cos^2 ny + q_{mn}^2 \cos^2(mx - \theta) \sin^2 ny$$
$$+ 2p_{mn}q_{mn} \cos(mx - \phi) \cos(mx - \theta) \cos ny \sin ny] \, dx \, dy.$$

Trigonometric reduction yields for the expression in brackets:

$$\tfrac{1}{4}(p_{mn}^2 + q_{mn}^2) + \tfrac{1}{4}p_{mn}^2[\cos 2(mx - \phi) + \cos 2ny + \cos 2(mx - \phi) \cos 2ny]$$
$$+ \tfrac{1}{4}q_{mn}^2[\cos 2(mx - \theta) - \cos 2ny - \cos 2(mx - \theta) \cos 2ny]$$
$$+ \tfrac{1}{2}p_{mn}q_{mn} \sin 2ny \, [\cos(2mx - \phi - \theta) + \cos(\theta - \phi)],$$

hence

$$I_{mn} = \tfrac{1}{4} \mid E_1 \mid + I'_{mn},$$

where I'_{mn} is a linear combination of eight Fourier coefficients of the periodic bounded function defined for $0 \leq x < 2\pi$, $0 \leq y < 2\pi$ by

$$f(x, y) = \begin{cases} 1 \text{ in } E_1, \\ 0 \text{ otherwise,} \end{cases}$$

the multipliers in the combination not exceeding π^2 numerically. Consequently, as m or n or both become infinite

$$I_{mn} \to \tfrac{1}{4} \mid E_1 \mid > 0,$$

from which our theorem follows.

3. **Generalization of Fatou-Denjoy-Lusin theorems to double series.**

THEOREM II. *Let $A_{mn}(x, y)$ be as defined in* (2.1) *or in the equivalent form* (2.8). *If the general double trigonometric series*

$$\sum_{m,n=1}^{\infty} A_{mn}(x, y)$$

is absolutely convergent in a two-dimensional point set E of positive measure $\mid E \mid >$ 0, then

$$\sum_{m,n=1}^{\infty} \rho_{mn} < \infty.$$

Proof. We may again dismiss the terms for which $\rho_{mn} = 0$. We define $Q(x, y)$ in E by

$$Q(x, y) = \sum_{\mu,\nu=1}^{\infty} \mid A_{\mu\nu}(x, y) \mid.$$

By assumption $Q(x, y)$ has a finite value for every (x, y) in E and is the limit of the sequence of continuous functions

$$Q_{mn}(x, y) = \sum_{\mu,\nu=1}^{m,n} | A_{\mu\nu}(x, y) |, \qquad (x, y) \text{ in } E.$$

An application of Egoroff's theorem asserts that to any $\epsilon > 0$ there exists a set E_1 contained in E with $| E_1 | > | E | - \epsilon > 0$, such that the sequence $Q_{mn}(x, y)$ converges uniformly in E_1 to $Q(x, y)$ which in E_1 has the finite bound B. In consequence of the uniform convergence we may integrate the series termwise over E_1 and get

$$\sum_{\mu,\nu=1}^{\infty} \iint_{E_1} | A_{\mu\nu}(x, y) | \, dx \, dy < B \, | E_1 |,$$

or, by (2.4),

$$\sum_{\mu,\nu=1}^{\infty} \rho_{\mu\nu} \iint_{E_1} | B_{\mu\nu}(x, y) | \, dx \, dy < B \, | E_1 |.$$

To prove our theorem it is sufficient to show that

$$J_{mn} \equiv \iint_{E_1} | B_{mn}(x, y) | \, dx \, dy$$

is bounded away from zero for all m, n; and in view of (2.5) it is sufficient to prove the same for I_{mn}. This follows (using the argument of §2) from $I_{mn} > 0$ and $I'_{mn} \to 0$ as one or both indices tend to ∞.

4. **Absolute convergence of double Fourier series.** For any function $f(x, y)$ we employ the following notation

$$\begin{aligned} \Delta_{11} f(x, y; s, t) \equiv \Delta_{11} f \equiv f(x + s, y + t) - f(x - s, y + t) \\ - f(x + s, y - t) + f(x - s, y - t), \end{aligned} \tag{4.1}$$

$$\Delta_{10} f(x, y; s) \equiv f(x + s, y) - f(x - s, y), \tag{4.2}$$

$$\Delta_{01} f(x, y; t) \equiv f(x, y + t) - f(x, y - t), \tag{4.3}$$

$$M_p(\Delta_{11} f) \equiv \left(\frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} | \Delta_{11} f(x, y; s, t) |^p \, dx \, dy \right)^{1/p}, \tag{4.4}$$

$$M_p(\Delta_{10} f) \equiv \left(\frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} | f(x + s, y) - f(x - s, y) |^p \, dx \, dy \right)^{1/p}, \tag{4.5}$$

$$M_p(\Delta_{01} f) \equiv \left(\frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} | f(x, y + t) - f(x, y - t) |^p \, dx \, dy \right)^{1/p}, \tag{4.6}$$

$$\lambda_{00} = \tfrac{1}{4}, \qquad \lambda_{m0} = \lambda_{0n} = \tfrac{1}{2}, \qquad \lambda_{mn} = 1 \qquad (m, n > 0),$$

$$(4.7)\qquad \begin{aligned} a_{mn} + ib_{mn} &= \frac{\lambda_{mn}}{\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f(u, v)\, e^{imu} \cos nv\, du\, dv \\ c_{mn} + id_{mn} &= \frac{\lambda_{mn}}{\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f(u, v)\, e^{imu} \sin nv\, du\, dv \end{aligned} \qquad (m, n \geq 0).$$

Thus with the notation of §2 and the coefficients (4.7) the double Fourier series of $f(x, y)$ is

$$(4.8)\qquad f(x, y) \sim \sum_{m,n=0}^{\infty} A_{mn}(x, y)$$

and it follows from (4.1) that if $f(x, y)$ is a function of one variable, then $\Delta_{11}f(x, y; s, t) \equiv 0$.

THEOREM III. *Suppose $f(x, y)$ is periodic and in L^2. Let $U_i(t)$ $(i = 1, 2)$ be positive non-increasing functions of t such that $U_i(t) \downarrow 0$ as $t \downarrow 0$. Let $V(s, t)$ be a positive function which is non-increasing as $s \downarrow 0$ or as $t \downarrow 0$. Suppose $0 < k \leq 2$ and*

$$(4.9)\qquad \sum_{\mu,\nu=1}^{\infty} \mu^{-\frac{1}{2}k}\nu^{-\frac{1}{2}k}\left[V\left(\frac{\pi}{\mu}, \frac{\pi}{\nu}\right)\right]^k < \infty,$$

$$(4.10)\qquad \sum_{\mu=1}^{\infty} \mu^{-\frac{1}{2}k}\left[U_i\left(\frac{\pi}{\mu}\right)\right]^k < \infty \qquad (i = 1, 2).$$

If $f(x, y)$ satisfies for $0 < s < \pi$, $0 < t < \pi$ each of the inequalities

$$(4.11)\qquad M_2(\Delta_{11}f) \leq V(s, t),$$

$$(4.12)\qquad M_2(\Delta_{10}f) \leq U_1(s),$$

$$(4.13)\qquad M_2(\Delta_{01}f) \leq U_2(t),$$

then

$$\sum_{\mu,\nu=0}^{\infty} \rho_{\mu\nu}^{k} < \infty.$$

Proof. With the aid of (2.8) and (4.8) we get

$$(4.14)\qquad \Delta_{11}f(x, y; s, t) \sim 4 \sum_{m,n=1}^{\infty} \sin ms \sin nt\, L_{mn}(x, y),$$

where

$$(4.15)\qquad L_{mn}(x, y) = p_{mn} \sin (mx - \phi) \sin ny - q_{mn} \sin (mx - \theta) \cos ny.$$

Write

$$(4.16)\qquad J_{mn} = \frac{1}{\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} L_{mn}^2(x, y)\, dx\, dy;$$

then from (4.15),

$$J_{mn} = p_{mn}^2 + q_{mn}^2 = \rho_{mn}^2 .$$

Now (4.14) is a Fourier series of a function of x and y in L^2; hence squaring (4.14) and integrating over $(-\pi, -\pi; \pi, \pi)$ we obtain

$$\begin{aligned} M_2^2(\Delta_{11}f) &\equiv \frac{1}{\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \mid \Delta_{11}f(x; y; s, t) \mid^2 dx\, dy \\ &= 16 \sum_{m,n=1}^{\infty} J_{mn} \sin^2 ms \sin^2 nt \\ &= 16 \sum_{m,n=1}^{\infty} \rho_{mn}^2 \sin^2 ms \sin^2 nt. \end{aligned}$$

By assumption (4.11) this gives

$$\sum_{m,n=1}^{\infty} \rho_{mn}^2 \sin^2 ms \sin^2 nt \le V^2(s, t).$$

Hence, a fortiori

$$\sum_{\mu,\nu=1}^{m,n} \rho_{\mu\nu}^2 \sin^2 \mu s \sin^2 \nu t \le V^2(s, t).$$

Now choose m, n arbitrarily and choose s and t such that

$$s = \frac{\pi}{2m}, \qquad t = \frac{\pi}{2n};$$

then using

$$\sin^2 \mu s > \frac{4}{\pi^2}\mu^2 s^2, \qquad \sin^2 \nu t > \frac{4}{\pi^2}\nu_n^2 t^2,$$

we get

$$\sum_{\mu,\nu=1}^{m-1,n-1} \rho_{\mu\nu}^2 \mu^2 \nu^2 \le m^2 n^2 V^2\left(\frac{\pi}{2m}, \frac{\pi}{2n}\right).$$

On replacing m by $2m$ and n by $2n$ we easily obtain

$$\sum_{\mu=m}^{2m-1}\sum_{\nu=n}^{2n-1} \rho_{\mu\nu}^2 \le 4V^2\left(\frac{\pi}{4m}, \frac{\pi}{4n}\right). \tag{4.17}$$

Now by Hölder's inequality for double sums [3],

$$\sum_{\mu=m}^{2m-1}\sum_{\nu=n}^{2n-1} \rho_{\mu\nu}^k \le m^{1/q} n^{1/q} \left(\sum_{\mu=m}^{2m-1}\sum_{\nu=n}^{2n-1} \rho_{\mu\nu}^2\right)^{\frac{1}{2}k} \qquad \left(q = \frac{2}{2-k} > 1\right),$$

and using (4.17),

$$\sum_{\mu=m}^{2m-1}\sum_{\nu=n}^{2n-1}\rho_{\mu\nu}^k \le 2^k m^{1-\frac{1}{2}k} n^{1-\frac{1}{2}k}\left[V\left(\frac{\pi}{4m},\frac{\pi}{4n}\right)\right]^k.$$

Since $V(s, t)$ does not increase as $s \downarrow 0$ or as $t \downarrow 0$, it follows that

$$\sum_{\mu=m}^{2m-1}\sum_{\nu=n}^{2n-1}\rho_{\mu\nu}^k \le 4^k \sum_{\mu=m}^{2m-1}\sum_{\nu=n}^{2n-1}\mu^{-\frac{1}{2}k}\nu^{-\frac{1}{2}k}\left[V\left(\frac{\pi}{\mu},\frac{\pi}{\nu}\right)\right]^k.$$

Take

$$m = 2^\eta, \qquad n = 2^\lambda \qquad (\eta = 0, 1, 2, \cdots; \lambda = 0, 1, 2, \cdots),$$

and sum over η and λ. This yields

$$\sum_{\mu,\nu=1}^{\infty}\rho_{\mu\nu}^k \le 4^k \sum_{\mu,\nu=1}^{\infty}\mu^{-\frac{1}{2}k}\nu^{-\frac{1}{2}k}\left[V\left(\frac{\pi}{\mu},\frac{\pi}{\nu}\right)\right]^k,$$

which by (4.9) gives

$$\sum_{\mu,\nu=1}^{\infty}\rho_{\mu\nu}^k < \infty.$$

We now wish to prove

$$\sum_{\mu=1}^{\infty}\rho_{\mu 0}^k < \infty, \qquad \sum_{\nu=1}^{\infty}\rho_{0\nu}^k < \infty.$$

We introduce the functions

$$g(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x,y)\,dy - \tfrac{1}{4}a_{00}, \qquad h(y) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x, y)\,dx - \tfrac{1}{4}a_{00}, \tag{4.18}$$

so that $g(x)$ and $h(y)$ have the Fourier series

$$\begin{aligned} g(x) &\sim \sum_{m=1}^{\infty} A_{m0}(x) = \sum_{m=1}^{\infty}(a_{m0}\cos mx + b_{m0}\sin mx),\\ h(y) &\sim \sum_{n=1}^{\infty} A_{0n}(y) = \sum_{n=1}^{\infty}(a_{0n}\cos ny + c_{0n}\sin ny). \end{aligned} \tag{4.19}$$

We now apply the Theorem 3.1 given by Szász [5; 376], where we put $m = 1$, $p = 2$ and obtain the condition

$$\sum_{n=1}^{\infty} n^{-\frac{1}{2}k}\,[M_2(\Delta g)]^k < \infty, \tag{4.20}$$

where

$$M_2(\Delta g) \equiv \left(\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|g\left(x+\frac{\pi}{n}\right) - g\left(x - \frac{\pi}{n}\right)\right|^2 dx\right)^{\frac{1}{2}}.$$

However, it is easily shown by the use of (4.18), Schwarz's inequality, and assumption (4.12) that

$$M_2(\Delta g(s)) \leq U_1(s). \tag{4.21}$$

Now from (4.10) we have

$$\sum_{n=1}^{\infty} n^{-\frac{1}{2}k} \left[U_1\left(\frac{\pi}{n}\right) \right]^k < \infty,$$

which in view of (4.21), with $s = \pi/n$, gives

$$\sum_{n=1}^{\infty} n^{-\frac{1}{2}k} \left[M_2\left(\Delta g\left(\frac{\pi}{n}\right)\right) \right]^k < \infty,$$

so that (4.20) is satisfied and Theorem 3.1 together with Remark 3.1 given by Szász [5; 376-378], when applied to the Fourier series of $g(x)$, gives

$$\sum_{\mu=1}^{\infty} \rho_{\mu 0}^{k} < \infty.$$

A similar argument from (4.13) and (4.10) and the application of Szász's Theorem 3.1 to the Fourier series of $h(y)$ gives

$$\sum_{\nu=1}^{\infty} \rho_{0\nu}^{k} < \infty.$$

Hence,

$$\sum_{\mu,\nu=0}^{\infty} \rho_{\mu\nu}^{k} < \infty,$$

and our theorem is proved.

For $k = 2$ (4.9) yields

$$\sum_{\mu,\nu=1}^{\infty} \rho_{\mu\nu}^{2} \log(\mu + 1) \log(\nu + 1) < \infty.$$

This follows from

$$\sum_{\mu,\nu=1}^{\infty} \rho_{\mu\nu}^{2} \int_0^{\pi} \frac{\sin^2 \mu s}{s}\, ds \int_0^{\pi} \frac{\sin^2 \nu t}{t}\, dt < \infty.$$

A similar remark holds for the $\rho_{0\nu}$, $\rho_{\mu 0}$.

The following theorem is a generalization of Theorem III.

THEOREM IV. *Let p, q, σ_1, σ_2 be constants such that*

$$\begin{gathered} 0 < p \leq 2 \leq q, \qquad 2 = \sigma_1 p + \sigma_2 q, \\ \sigma_1 + \sigma_2 = 1, \qquad \sigma_1 \geq 0, \qquad \sigma_2 \geq 0. \end{gathered} \tag{4.22}$$

Assume $f(x, y)$ is in L^q and assume $V_1(s, t)$ and $V_2(s, t)$ are functions which decrease as $s \downarrow 0$ or as $t \downarrow 0$ and such that

$$\sum_{\mu=1}^{\infty} \sum_{\nu=1}^{\infty} \mu^{-\frac{1}{2}k} \nu^{-\frac{1}{2}k} \left\{ \left[V_1\left(\frac{\pi}{\mu}, \frac{\pi}{\nu}\right) \right]^{\sigma_1 p} \left[V_2\left(\frac{\pi}{\mu}, \frac{\pi}{\nu}\right) \right]^{\sigma_2 q} \right\}^{\frac{1}{2}k} < \infty, \ 0 < k \leq 2. \tag{4.23}$$

If $f(x, y)$ satisfies each of the inequalities (4.12), (4.13) *and*

$$M_p(\Delta_{11} f) \leq V_1(s, t), \qquad M_q(\Delta_{11} f) \leq V_2(s, t), \tag{4.24}$$

then

$$\sum_{\mu, \nu=0}^{\infty} \rho_{\mu\nu}^k < \infty.$$

Proof. We first prove

$$M_2^2(\Delta_{11} f) \leq M_p^{\sigma_1 p}(\Delta_{11} f)\, M_q^{\sigma_2 q}(\Delta_{11} f). \tag{4.25}$$

We introduce r_1, r_2 such that $r_1 + r_2 = 2$, $r_1 \geq 0$, $r_2 \geq 0$. Now

$$M_2^2(\Delta_{11} f) = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} |\Delta_{11} f|^{r_1} \, |\Delta_{11} f|^{r_2} \, dx\, dy.$$

An application of Hölder's inequality for double integrals gives

$$J_2 \equiv M_2^2(\Delta_{11} f) \leq J_{r_1\lambda}^{1/\lambda} J_{r_2\lambda'}^{1/\lambda'} = M_{r_1\lambda}^{r_1}(\Delta_{11} f)\, M_{r_2\lambda'}^{r_2}(\Delta_{11} f), \tag{4.26}$$

where $\lambda > 1$ and $1/\lambda + 1/\lambda' = 1$. Now we choose r_1, r_2 and λ such that

$$r_1 = \sigma_1 p, \qquad r_2 = \sigma_2 q, \qquad r_1 \lambda = p.$$

Then our conditions on r_1, r_2 are satisfied. Now (4.26) becomes

$$M_2^2(\Delta_{11} f) \leq M_p^{\sigma_1 p}(\Delta_{11} f)\, M_q^{\sigma_2 q}(\Delta_{11} f),$$

which is the desired relation (4.25). On using (4.24) we get from (4.25) that

$$M_2^2(\Delta_{11} f) \leq [V_1(s, t)]^{\sigma_1 p} \, [V_2(s, t)]^{\sigma_2 q}.$$

If we now define

$$V(s, t) = [V_1(s, t)^{\sigma_1 p} \, V_2(s, t)^{\sigma_2 q}]^{\frac{1}{2}},$$

then, by (4.23), it is seen that (4.9) is satisfied. Thus all assumptions of Theorem III are satisfied and Theorem IV follows from Theorem III.

Application of Theorem IV. We choose

$$V_1(s, t) = Cs^{\alpha_1} t^{\beta_1}, \qquad V_2(s, t) = Cs^{\alpha_2} t^{\beta_2}. \tag{4.27}$$

We also note that by (4.22) only two of the four parameters p, q, σ_1, σ_2 are independent. Therefore, we write

$$\sigma_1 = \sigma, \qquad \sigma_2 = 1 - \sigma,$$

and from

$$2 = \sigma p + (1 - \sigma)q = q - (q - p)\sigma$$

we get

$$\sigma = \frac{q-2}{q-p}, \qquad \lambda = \frac{1}{\sigma_1} = \frac{q-p}{q-2},$$

and correspondingly

$$\sigma_1 p = r_1 = p\,\frac{q-2}{q-p}, \qquad \sigma_2 q = r_2 = q\,\frac{2-p}{q-p}.$$

Now assumption (4.23) becomes

$$\sum_{\mu=1}^{\infty}\sum_{\nu=1}^{\infty} \mu^{-\frac{1}{2}k}\nu^{-\frac{1}{2}k}\left\{\left[\mu^{-\alpha_1}\nu^{-\beta_1}\right]^{p(q-2)/(q-p)}\left[\mu^{-\alpha_2}\nu^{-\beta_2}\right]^{q(2-p)/(q-p)}\right\}^{\frac{1}{2}k} < \infty, \tag{4.28}$$

which is satisfied whenever

$$\frac{k}{2}\left[1 + \frac{\alpha_1 p(q-2) + \alpha_2 q(2-p)}{q-p}\right] > 1$$

and

$$\frac{k}{2}\left[1 + \frac{\beta_1 p(q-2) + \beta_2 q(2-p)}{q-p}\right] > 1,$$

or

$$\begin{aligned} &k[q - p + pq(\alpha_1 - \alpha_2) - 2\alpha_1 p + 2\alpha_2 q] > 2(q - p),\\ &k[q - p + pq(\beta_1 - \beta_2) - 2\beta_1 p + 2\beta_2 q] > 2(q - p). \end{aligned} \tag{4.29}$$

Assumption (4.24) becomes

$$M_p(\Delta_{11} f) \le C s^{\alpha_1} t^{\beta_1}, \qquad M_q(\Delta_{11} f) \le C s^{\alpha_2} t^{\beta_2}. \tag{4.30}$$

Now choose $\alpha_1 = \beta_1 = p = 1$ and write $\alpha_2 = \alpha$, $\beta_2 = \beta$; then for this special case (4.29) becomes

$$k[q(2+\alpha) - 3] > 2(q-1), \qquad k[q(2+\beta) - 3] > 2(q-1), \tag{4.31}$$

and (4.30) becomes

$$M_1(\Delta_{11} f) \le Cst, \qquad M_q(\Delta_{11} f) \le C s^{\alpha} t^{\beta}. \tag{4.32}$$

We now assume that $f(x, y)$ satisfies a Lipschitz condition Lip (α, β) defined by

$$\begin{aligned} &| f(x+s, y+t) - f(x-s, y+t) - f(x+s, y-t) + f(x-s, y-t) | \\ &\qquad \le C s^{\alpha} t^{\beta}, \end{aligned} \tag{4.33}$$

$\alpha > 0$, $\beta > 0$, and that $f(x, y)$ is of bounded variation H [4; I, 345]. We shall

prove that functions of bounded variation H satisfy the integrated Lipschitz condition given in the first condition in (4.32). Also (4.33) implies that the second condition in (4.32) is satisfied for every $q > 0$. Thus we can take q arbitrarily large, and so, on allowing $q \to \infty$, we find that (4.31) and (4.32) reduce to

$$k > \frac{2}{2 + \alpha}, \qquad k > \frac{2}{2 + \beta}, \tag{4.34}$$

and

$$M_1(\Delta_{11}f) \leq Cst, \qquad |\Delta_{11}f| \leq Cs^\alpha t^\beta. \tag{4.35}$$

We now give a proof of the result used above that functions of bounded variation H satisfy the Lipschitz condition [2; 37, Theorem VI]

$$M_1(\Delta_{11}f) \leq Cst$$

in the square $-\pi \leq s, t \leq \pi$. We write

$$M_1(\Delta_{11}f) = \frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} | f(x + s, y + t) - f(x - s, y + t)$$
$$-f(x + s, y - t) + f(x - s, y - t) \mid dx\, dy.$$

Now using the concept of a Stieltjes integral we have

$$f(x + s, y + t) - f(x - s, y + t) = \int_{x-s}^{x+s} d_x f(x, y + t),$$

if $f(x, y)$ is of bounded variation in x for constant $y + t$;

also

$$f(x + s, y - t) - f(x - s, y - t) = \int_{x-s}^{x+s} d_x f(x, y - t);$$

hence

$$M_1(\Delta_{11}f) = \frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\left|\int_{x-s}^{x+s} d_u\{f(u, y + t) - f(u, y - t)\}\right| dx\, dy.$$

Let $f(x, y)$ be of bounded variation in y for every x; then

$$f(u, y + t) - f(u, y - t) = \int_{y-t}^{y+t} d_v f(u, v);$$

hence

$$4\pi^2 M_1(\Delta_{11}f) = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} dx\, dy\left|\int_{x-s}^{x+s} d_u \int_{y-t}^{y+t} d_v f(u, v)\right|.$$

Now

$$| \Delta_{11} f | = \left| \int_{x-s}^{x+s} d_u \{ f(u, y+t) - f(u, y-t) \} \right|$$

$$\leq \int_{x-s}^{x+s} | d_u \{ f(u, y+t) - f(u, y-t) \} |,$$

so that

$$4\pi^2 M_1(\Delta_{11} f) \leq \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} dx\, dy \int_{x-s}^{x+s} \left| d_u \int_{y-t}^{y+t} | d_v f(u, v) | \right|.$$

As

$$x - s \leq u \leq x + s, \qquad y - t \leq v \leq y + t,$$

we have

$$4\pi^2 M_1(\Delta_{11} f) \leq \left(\int_{-\pi-s}^{\pi+s} \left| d_u \int_{-\pi-t}^{\pi+t} | d_v f(u, v) | \right| \right) \left(\int_{y-t}^{y+t} dv \right) \left(\int_{x-s}^{x+s} du \right).$$

Suppose $f(x, y)$ has a finite total variation

$$V(f) = \text{l.u.b.} \sum_{\mu, \nu = 0}^{m-1, n-1} | f(u_{\mu+1}, v_{\nu+1}) - f(u_{\mu+1}, v_\nu) - f(u_\mu, v_{\nu+1}) + f(u_\mu, v_\nu) |;$$

then the above inequality yields

$$M_1(\Delta_{11} f) \leq stV(f)$$

since

$$\int_{-\pi}^{\pi} | d_v f(u, v) | = \text{l.u.b.} \sum_{\nu=0}^{n-1} | f(u, v_{\nu+1}) - f(u, v_\nu) | = F(u),$$

$$\int_{-\pi}^{\pi} | d_u F(u) | = \text{l.u.b.} \sum_{\mu=0}^{m-1} | F(u_{\mu+1}) - F(u_\mu) |$$

$$= \text{l.u.b.} \sum_{\mu=0}^{m-1} \left| \sum_{\nu=0}^{n-1} | f(u_{\mu+1}, v_{\nu+1}) - f(u_{\mu+1}, v_\nu) | - \sum_{\nu=0}^{n-1} | f(u_\mu, v_{\nu+1}) - f(u_\mu, v_\nu) | \right| < V(f).$$

Since a function of bounded variation H has $V(f)$ finite, is of bounded variation in x for every y and is of bounded variation in y for every x, our proof is completed.

Summarizing we have the

COROLLARY. *Suppose that $f(x, y)$ satisfies the Lipschitz condition* (4.33) *for some $\alpha > 0$, $\beta > 0$, and that it is of bounded variation H; let*

$$k > \max\left(\frac{2}{2+\alpha}, \frac{2}{2+\beta}\right);$$

then

$$\sum_{\mu,\nu=0}^{\infty} \rho_{\mu\nu}^{k} < \infty.$$

Note. To construct a counter example for the purpose that $k = 2/(2 + \alpha)$, $\alpha = \beta$, no longer gives absolute convergence of the Fourier coefficients, consult Zygmund's function $z(x)$ in [6; 591–598] and in [7; 138] and consider the double Fourier series of $f(x, y) = z(x)z(y)$.

For a counter example for Theorem III we may either consider special functions V (like the Lipschitz condition), or try to copy Szász's method in his Transactions paper and use functions of the type $f(x, y) = z(x)u(y)$.

BIBLIOGRAPHY

1. P. FATOU, *Sur la convergence absolue des séries trigonométriques*, Bulletin de la Société Mathématique de France, vol. 41(1913), pp. 47–53.
2. J. J. GERGEN, *Convergence criteria for double Fourier series*, Transactions of the American Mathematical Society, vol. 35(1933), pp. 29–63.
3. G. H. HARDY, J. E. LITTLEWOOD, AND G. PÓLYA, *Inequalities*, Cambridge, 1934.
4. E. W. HOBSON, *Theory of Functions of a Real Variable*, Cambridge, 1927.
5. O. SZÁSZ, *Fourier series and mean moduli of continuity*, Transactions of the American Mathematical Society, vol. 42(1937), pp. 366–395.
6. A. ZYGMUND, *Some points in the theory of trigonometric and power series*, Transactions of the American Mathematical Society, vol. 36(1934), pp. 586–617.
7. A. ZYGMUND, *Trigonometrical Series*, Warsaw-Lwów, 1935.

INTEGRAL FORMULAS IN CROFTON'S STYLE ON THE SPHERE AND SOME INEQUALITIES REFERRING TO SPHERICAL CURVES

By L. A. Santaló

Introduction. Several integral formulas referring to convex plane curves, notable for their great generality, were obtained by W. Crofton in 1868 and successive years from the theory of geometrical probability [6], [7], [8], [9], [10].

A direct and rigorous exposition of Crofton's principal results, adding some new formulas, was made in 1912 by H. Lebesgue [12]. Another systematic exposition of Crofton's most interesting formulas, together with the generalization of many of them to space, is found in the two volumes on integral geometry by Blaschke [2].

The purpose of the present paper is to give a generalization of Crofton's formulas to the surface of the sphere. This is what we do in part I. We find further integral formulas on the sphere (for instance, (16), (17), (20), (21)) which have no equivalent in the plane. Other formulas, if we consider the plane as the limit of a sphere whose radius increases indefinitely, give integral formulas referring to plane convex curves (e. g., (34), (35)) which we think are new.

In part II, with simple methods of integral geometry [2], we obtain three inequalities referring to spherical curves. Inequality (38) is the generalization to the sphere of an inequality that Hornich [11] obtained for plane curves. (52) and (58) contain the classical isoperimetric inequality on the sphere. Finally, inequality (61) gives a superior limitation for the "isoperimetric deficit" of convex curves on the sphere.

I. FORMULAS IN THE CROFTON STYLE ON THE SPHERE

1. **Notation and useful formulas.** The element of area on the sphere of unit radius will be represented by $d\Omega$; that is, if θ and φ are the spherical coördinates of the point Ω, we have

$$d\Omega = \sin\theta \, d\theta \, d\varphi. \tag{1}$$

A great non-directed circle C on the same sphere of unit radius can be determined by one of its poles, that is, by either of the extremities of the diameter perpendicular to it. Since $d\Omega$ is the element of area of one of these extremities, the "density" for measuring sets of great circles on the sphere is [2; 61, 80]

$$dC = d\Omega; \tag{2}$$

Received January 9, 1942.

that is, the "measure" of a set of great circles on the sphere is defined as the integral of (2) extended over this set.

It is possible to give the density (2) another form, which will sometimes be useful. We consider a fixed great circle C_0 and a fixed point A on it. The great circle C can be determined for the abscissa t of one of the intersection points from C and C_0 and the angle α between the two circles. If θ and φ are the spherical coördinates of the pole Ω of C with regard to the pole Ω_0 of C_0, $\theta = \alpha$, $\varphi = t$, and (1), (2) give

$$dC = \sin \alpha \, d\alpha \, dt. \tag{3}$$

Let us consider two great circles C_1, C_2 and one of their intersection points Ω. If α_1 and α_2 are the angles that C_1 and C_2 make with another fixed great circle which also passes through Ω, the following differential formula [2; 78] is known:

$$dC_1 \, dC_2 = | \sin (\alpha_2 - \alpha_1) | \, d\alpha_1 \, d\alpha_2 \, d\Omega. \tag{4}$$

By (2), formula (4) can be transformed into a "dual" form. Let Ω_1 and Ω_2 be two points on the unit sphere and let C be the great circle determined by them. If β_1 and β_2 are the abscissas of Ω_1 and Ω_2 on C in relation to a fixed origin on this circle, (4) is equivalent to

$$d\Omega_1 \, d\Omega_2 = | \sin (\beta_1 - \beta_2) | \, d\beta_1 \, d\beta_2 \, dC. \tag{5}$$

2. **First integral formulas. Convex curves on the sphere.** A closed curve on the sphere is said to be *convex* when it cannot be cut by a great circle in more than two points.

A convex curve divides the surface of the sphere into two parts, one of which is always wholly contained in a hemisphere; that is, there is always a great circle which has the whole convex curve on the same side; we only have to consider, for example, a tangent great circle.

When we say a "convex figure", we understand that part of the surface of the sphere which is limited by a convex curve and is smaller than or equal to a hemisphere.

Let us consider a convex figure K on the sphere of unit radius. The radii which are perpendicular to the tangent planes (or, more generally, to the planes of support) to the cone which projects K from the center of the sphere form another cone whose intersection with the sphere is a new convex curve K^*. We shall call K^* the "dual" curve of K. The lengths and areas of K and K^* are connected by the known relations

$$F^* = 2\pi - L, \qquad L^* = 2\pi - F. \tag{6}$$

All the great circles C that cut K have their poles in the area bounded by K^* and the symmetrical curve of the same K^* with respect to the center of the

sphere. This area equals $4\pi - 2F^* = 2L$. Counting each pair of points which are the extremities of a diameter as a single point, and taking into account the value (2) of the density dC, we have

$$\int_{C \cdot K \neq 0} dC = L; \tag{7}$$

this means: on the sphere, the measure of the great circles which cut a convex curve is equal to the length of this curve. This result is given by [2; 81].

3. **Integral of the chords.** Let Ω_1 and Ω_2 be two points inside the convex curve K (always on the unit sphere) and let C be the great circle determined by them. The differential expression (5) can be integrated for all pairs of points within K.

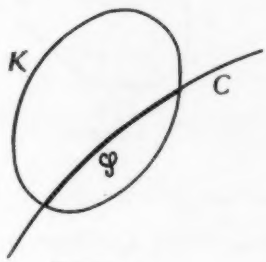

FIGURE 1

The integral of the left side is F^2. By calculating the integral of the right side, if φ represents the length of the arc of C that is contained in K (Fig. 1), we have

$$\int_0^{\varphi} \int_0^{\varphi} |\sin(\beta_1 - \beta_2)| \, d\beta_1 \, d\beta_2 = 2(\varphi - \sin\varphi). \tag{8}$$

Hence

$$\int_{C \cdot K \neq 0} (\varphi - \sin\varphi) \, dC = \tfrac{1}{2}F^2. \tag{9}$$

This formula generalizes, as we shall see (§11), Crofton's formula for chords in plane geometry.

4. **Principal Crofton formula.** Let us consider all the pairs of great circles C_1, C_2 that cut K. From (7) we deduce

$$\int_{\substack{C_1 \cdot K \neq 0 \\ C_2 \cdot K \neq 0}} dC_1 \, dC_2 = L^2. \tag{10}$$

Now we can make the integration of formula (4) extend only to the pairs of great circles which cut K. If Ω is fixed inside K, α_1 and α_2 can vary from 0 to π and

$$(11)\qquad \int_{\Omega\subset K}\left(\int_0^\pi\int_0^\pi |\sin(\alpha_1-\alpha_2)|\,d\alpha_1\,d\alpha_2\right)d\Omega = 2\pi\int_{\Omega\subset K} d\Omega = 2\pi F;$$

if Ω is outside K, α_1 and α_2 can vary from 0 to the angle ω between the great circles which are tangent to K and which pass through Ω (Fig. 2). By applying

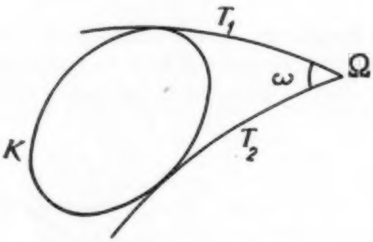

FIGURE 2

(8), the value of this last integral is found to be $\int 2(\omega - \sin\omega)\,d\Omega$ for $\Omega \not\subset K$. Adding this result to (11), we have (10); hence

$$(12)\qquad \int(\omega - \sin\omega)\,d\Omega = \tfrac{1}{2}L^2 - \pi F \qquad (\Omega \not\subset K).$$

This formula has the same form as Crofton's fundamental formula of plane geometry. The integration in (12) is extended to all points Ω outside K, each pair of points situated in the extremities of a diameter being considered as a single point.

5. **"Dual" formulas.** From a convex curve K we can deduce the "dual" curve K^* as we have seen in §2. To a great circle C which cuts K corresponds a point Ω^* (the pole of C) which is not inside K^*. The arc φ of C inside K is equal to $\pi - \omega^*$, ω^* being the angle between the two great circles tangent to K^* drawn through Ω^*. Since $F = 2\pi - L^*$ (by (6)), formula (9) can be written

$$(13)\qquad \int(\pi - \omega^* - \sin\omega^*)\,d\Omega^* = \tfrac{1}{2}(2\pi - L^*)^2 \qquad (\Omega^* \not\subset K^*).$$

The integration is extended over the outside of K^* (the points which are the extremities of the same diameter being considered as a single point) and consequently

$$\int \pi\,d\Omega^* = \pi(2\pi - F^*) \qquad (\Omega^* \not\subset K^*).$$

Then (13) gives

$$\int (\omega^* + \sin \omega^*)\, d\Omega^* = 2\pi L^* - \pi F^* - \tfrac{1}{2}L^{*2} \qquad (\Omega^* \not\subset K^*). \tag{14}$$

This formula holds for any convex curve K^*; hence it is valid for K:

$$\int (\omega + \sin \omega)\, d\Omega = 2\pi L - \pi F - \tfrac{1}{2}L^2 \qquad (\Omega \not\subset K). \tag{15}$$

From (15) and (12), we deduce

$$\int \omega\, d\Omega = \pi L - \pi F \qquad (\Omega \not\subset K) \tag{16}$$

and

$$\int \sin \omega\, d\Omega = \pi L - \tfrac{1}{2}L^2 \qquad (\Omega \not\subset K). \tag{17}$$

The same procedure shows that formula (12) is equivalent to

$$\int_{C^* \cdot K^* \neq 0} (\pi - \varphi^* - \sin \varphi^*)\, dC^* = \tfrac{1}{2}(2\pi - F^*)^2 - \pi(2\pi - L^*), \tag{18}$$

where the integration is extended over all the great circles C^* which cut K^*. By (7) we have $\pi \int dC^* = \pi L^*$ and by substitution of this value in (18) and writing the formula for K, we have

$$\int_{C \cdot K \neq 0} (\varphi + \sin \varphi)\, dC = 2\pi F - \tfrac{1}{2}F^2, \tag{19}$$

where φ is the length of the arc of C which is inside K.

From (9) and (19) we deduce

$$\int_{C \cdot K \neq 0} \varphi\, dC = \pi F, \tag{20}$$

and

$$\int_{C \cdot K \neq 0} \sin \varphi\, dC = \pi F - \tfrac{1}{2}F^2. \tag{21}$$

We repeat. In (16), (17), ω is the angle between the two great circles tangent to K through Ω; in (20), (21), φ is the length of the arc of the great circle C which is inside K.

The formulas (16), (17), (20), (21) that hold for any convex curve on the unit sphere have no equivalent in the plane.

6. **Formulas for the tangents.** Let K be a convex curve on the unit sphere with continuous radius of geodesic curvature.

If τ is the angle between a variable tangent great circle and a fixed tangent great circle and if s is the length of the arc of K, the radius of geodesic curvature ρ_g is given by

$$\rho_g = \frac{ds}{d\tau}, \tag{22}$$

and the Gauss-Bonnet formula gives

$$\oint_K \frac{ds}{\rho_g} = \int d\tau = 2\pi - F. \tag{23}$$

Let us consider two great circles tangent to K; let Ω be one of the intersection points of these circles. T_1 and T_2 will be the lengths of the arcs of these great circles bounded by Ω and the points of contact (T_1 and $T_2 \leq \pi$), and we represent by ω the angle between the two tangent circles at Ω (Fig. 2).

We wish to express the element of area $d\Omega$ as a function of the angles τ_1, τ_2 which determine the tangent great circles.

For fixed τ_2, as we pass from τ_1 to $\tau_1 + d\tau_1$, the arc T_2 is increased by $dT_2 = (\sin T_1/\sin \omega)d\tau_1$.

In the same way, as we pass from τ_2 to $\tau_2 + d\tau_2$, the arc T_1 is increased by $dT_1 = (\sin T_2/\sin \omega)d\tau_2$.

Since the element of area $d\Omega$ can be expressed in the form $d\Omega = \sin \omega \, dT_1 \, dT_2$, we find the desired expression

$$d\Omega = \frac{\sin T_1 \cdot \sin T_2}{\sin \omega} d\tau_1 \, d\tau_2$$

or

$$\frac{\sin \omega}{\sin T_1 \cdot \sin T_2} d\Omega = d\tau_1 \, d\tau_2 . \tag{24}$$

7. We can make the integration of (24) extend over all pairs of circles tangent to K and, by counting each pair once only (to do this we must divide the integral by 2), we have, by (23),

$$\int \frac{\sin \omega}{\sin T_1 \cdot \sin T_2} d\Omega = \tfrac{1}{2}(2\pi - F)^2 \qquad (\Omega \not\subset K). \tag{25}$$

Likewise, as in the preceding cases, the notation $\Omega \not\subset K$ indicates that the integration must be extended over all points Ω outside K; the points situated in the extremities of a diameter are considered as a single point.

8. Let $\rho_g^{(1)}$, $\rho_g^{(2)}$ be the radii of geodesic curvature of K at the points of contact of the tangent great circles through Ω. By virtue of (22), (24), we have

$$\sin \omega \frac{\rho_g^{(1)} \rho_g^{(2)}}{\sin T_1 \cdot \sin T_2} \, d\Omega = ds_1 \, ds_2 .$$

By integrating this expression over all pairs of tangent great circles, counting each pair once only, we get

$$\int \sin \omega \frac{\rho_g^{(1)} \rho_g^{(2)}}{\sin T_1 \cdot \sin T_2} \, d\Omega = \tfrac{1}{2}L^2 \qquad (\Omega \not\subset K). \tag{26}$$

9. By (22) and (24), we have

$$\sin \omega \frac{\rho_g^{(1)}}{\sin T_1 \cdot \sin T_2} \, d\Omega = ds_1 \, d\tau_2$$

and by integrating over all great circles tangent to K and observing that each point Ω is a common factor of two terms, it follows that

$$\int \sin \omega \frac{\rho_g^{(1)} + \rho_g^{(2)}}{\sin T_1 \cdot \sin T_2} \, d\Omega = L(2\pi - F) \qquad (\Omega \not\subset K). \tag{27}$$

10. **"Dual" formulas.** According to §5, from formulas (25), (26), (27) we can deduce the respective "dual" formulas.

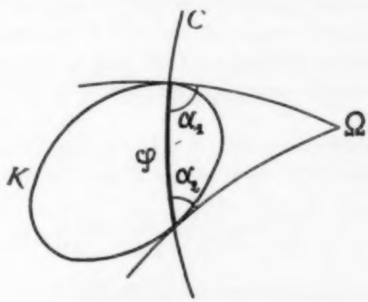

FIGURE 3

If φ is the length of the arc of the great circle C which is inside K and α_1, α_2 are the angles that C makes with the great circles tangent to K at the intersection points of C with K (Fig. 3), formula (25) gives

$$\int_{C \cdot K \neq 0} \frac{\sin \varphi}{\sin \alpha_1 \cdot \sin \alpha_2} \, dC = \tfrac{1}{2}L^2. \tag{28}$$

We observe that the "dual" element of ds is $d\tau^*$ for the dual curve K^* and

reciprocally. Then the dual expression of $\rho_s = ds/d\tau$ will be $d\tau^*/ds^* = 1/\rho_s^*$. Hence, formula (26) gives

$$\int_{C\cdot K\neq 0} \frac{\sin\varphi}{\sin\alpha_1\cdot\sin\alpha_2}\cdot\frac{1}{\rho_s^{(1)}\rho_s^{(2)}}\, dC = \tfrac{1}{2}(2\pi - F)^2. \tag{29}$$

Likewise, formula (27) gives

$$\int_{C\cdot K\neq 0} \frac{\sin\varphi}{\sin\alpha_1\cdot\sin\alpha_2}\left(\frac{1}{\rho_s^{(1)}}+\frac{1}{\rho_s^{(2)}}\right) dC = (2\pi - F)L. \tag{30}$$

11. **Passage to the case of the plane.** The classical Crofton formulas for the plane must result as a special case of the preceding formulas when the radius of the sphere increases indefinitely. Moreover, by this procedure, we shall find some new integral formulas.

We observe the following. (i) The element of area $d\Omega$ on the unit sphere can be replaced by dP/R^2, where dP is the element of area on the sphere of radius R and, as $R \to \infty$, dP will be the element of area in the plane. (ii) Let us consider the form (3) for dC; for the sphere of radius R this expression (3) must be replaced by $dC = \sin\alpha\, d\alpha(dt_R/R)$, where t_R is the length of the arc of the great circle of the sphere of radius R; when R increases to ∞, (3) is lim $R\cdot dC = dG$, dG being the "density" of the straight lines of the plane (recall that the "density" dG can be written $dG = \sin\alpha\, d\alpha\, dt$, where α is the angle which G forms with another fixed straight line and t is the abscissa of the intersection point [2; 7]). (iii) When we consider a sphere of radius R, the area F and length L which are in formulas from §§2-10 must be replaced by F/R^2 and L/R, respectively.

When these remarks are taken into account, the preceding formulas give the following results.

(i) Let us consider formula (9). If σ is the length of the arc that the great circle C determines in K, then $\varphi = \sigma/R$ and for R large we have

$$\varphi - \sin\varphi = \frac{\sigma^3}{3!R^3} - \frac{\sigma^5}{5!R^5} + \cdots .$$

If dC and F are replaced in (9) by dG/R and F/R^2, as $R \to \infty$ we have

$$\int_{G\cdot K\neq 0} \sigma^3\, dG = 3F^2. \tag{31}$$

This is the classical chord formula from Crofton [9; 84], [10; 27], [2; 20].

(ii) Formula (12) maintains the same form for the plane. Indeed, ω and $\sin\omega$ do not change; $d\Omega$ becomes dP/R^2, F becomes F/R^2, and L becomes L/R; in the limit as $R \to \infty$, formula (12) does not change. It is the "principal" Crofton formula for the plane [9; 78], [10; 26], [2; 18].

(iii) Formulas (16), (17), (20), (21) have no equivalent in the plane, since,

when these formulas are written for the sphere of radius R, as $R \to \infty$ the right side increases indefinitely.

(iv) In formula (25), we must replace $\sin T_1$ and $\sin T_2$ by T_1/R and T_2/R, the element of area $d\Omega$ by dP/R^2, and F by F/R^2. In the limit as $R \to \infty$, we find

$$\int \frac{\sin \omega}{T_1 T_2}\, dP = 2\pi^2 \qquad (P \not\subset K). \tag{32}$$

In this well-known formula ([12]; see also W. Blaschke, *Differentialgeometrie* I, p. 49), T_1 and T_2 are the lengths of the tangents to the convex curve K drawn through P, $dP = dx\, dy$ is the element of area on the plane, and ω is the angle between the tangents at P.

For formulas (26), (27), it is only necessary to observe that the radii of geodesic curvature become the radii of the ordinary curvature of the plane curve. Hence formulas (26) and (27) give the known formulas [12]

$$\int \sin \omega \frac{\rho_1 \rho_2}{T_1 T_2}\, dP = \tfrac{1}{2} L^2, \qquad \int \sin \omega \frac{\rho_1 + \rho_2}{T_1 T_2}\, dP = 2\pi L \tag{33}$$
$$(P \not\subset K).$$

(v) Formula (28), when φ is replaced by σ/R (σ is the length of the arc that the great circle C determines in K and in the limit it is the length of the chord that the straight line G determines in K) and R increases indefinitely, gives

$$\int_{G \cdot K \neq 0} \frac{\sigma}{\sin \alpha_1 \cdot \sin \alpha_2}\, dG = \tfrac{1}{2} L^2. \tag{34}$$

α_1 and α_2 are the angles that the straight line G makes with the tangents to K at the intersection points of G with K. The integration in (34) is extended over all the straight lines G which cut K.

Likewise, (29) and (30) give for the plane

$$\begin{aligned} &\int_{G \cdot K \neq 0} \frac{\sigma}{\rho_1 \rho_2 \sin \alpha_1 \sin \alpha_2}\, dG = 2\pi^2, \\ &\int_{G \cdot K \neq 0} \frac{(\rho_1 + \rho_2)\sigma}{\rho_1 \rho_2 \sin \alpha_1 \sin \alpha_2}\, dG = 2\pi L, \end{aligned} \tag{35}$$

where ρ_1 and ρ_2 are the radii of curvature of the convex curve K at the intersection points of G with K.

II. SOME INEQUALITIES REFERRING TO SPHERICAL CURVES

12. **A known formula.** Hitherto we have only considered relations on the sphere between a convex curve K and points and great circles. Now we wish to

establish some new relations which arise from considering on the sphere sets of variable small circles of constant spherical radius.

Let $\mathfrak{L}$ be a rectifiable curve (not necessarily convex) of length L on the unit sphere. We consider on the same sphere a small circle C_0 of spherical radius ρ ($\rho \leq \frac{1}{2}\pi$), whose length and area will be

$$L_0 = 2\pi \sin \rho, \qquad F_0 = 2\pi(1 - \cos \rho). \tag{36}$$

Let Ω be the center of the circle C_0 and, as in §1, $d\Omega$ the corresponding element of area of the sphere. If n represents the number of intersection points of the curve $\mathfrak{L}$ with the circle C_0 (n will be a function of Ω), we have the known formula

$$\int n \, d\Omega = \frac{2}{\pi} L L_0 \,,$$

or

$$\int n \, d\Omega = 4L \sin \rho; \tag{37}$$

the integration is extended over the whole sphere.

This formula is a particular case of Poincaré's formula of integral geometry [2; 81]. In [2], the formula is established only for spherical curves composed of a finite number of arcs with a continuously turning tangent. More generally, formula (37) is also valid for the case of a curve $\mathfrak{L}$ only supposed to be rectifiable and a circle C_0. The proof can be copied step by step from that given for Euclidean space of n dimensions in [13].

13. **An inequality referring to rectifiable curves on the sphere.** In this section we generalize for curves on the sphere an inequality that Hornich obtained for Euclidean space [11]. The proof is analogous to that given for Euclidean space in [13].

Let us consider on the sphere of unit radius the rectifiable curve $\mathfrak{L}$ of length L. Let F be the area filled by the points of the sphere whose spherical distance from $\mathfrak{L}$ is $\rho \leq \frac{1}{2}\pi$.

We shall prove that

$$F \leq 2L \sin \rho + 2\pi(1 - \cos \rho) \tag{38}$$

and establish the conditions for the equality in (38).

Let M_i ($i = 0, 1, 2, 3, \cdots$) be the area covered by the centers of the circles of radius ρ whose distance to $\mathfrak{L}$ is not greater than ρ and which have i points in common with $\mathfrak{L}$.

By (37), we have

$$M_1 + 2M_2 + 3M_3 + 4M_4 + \cdots = 4L \sin \rho, \tag{39}$$

and according to the definition of the area F,

$$M_0 + M_1 + M_2 + M_3 + \cdots = F. \tag{40}$$

From (39) and (40) we deduce

$$2F - 4L \sin \rho = 2M_0 + M_1 - (M_3 + 2M_4 + \cdots). \tag{41}$$

We consider the arc of a great circle of length D ($\leq \pi$) which joins the extremities of the given curve $\mathfrak{L}$ (if this curve is closed, $D = 0$). Let us call M_i^* ($i = 0, 1, 2$) the area covered by the centers of the circles of spherical radius ρ which have i points in common with this arc of length D (for $i = 0$ the arc is interior to the circle).

The area filled by the points whose distance from the arc of length D is not less than $\rho \leq \frac{1}{2}\pi$ is limited by two arcs of circles parallel to this arc at the distance ρ and two semicircles of radius ρ at the ends. The value of this area is $2D \sin \rho + 2\pi(1 - \cos \rho)$ and we can write

$$M_0^* + M_1^* + M_2^* = 2D \sin \rho + 2\pi(1 - \cos \rho). \tag{42}$$

By (37) we have also

$$M_1^* + 2M_2^* = 4D \sin \rho. \tag{43}$$

From (42) and (43) we deduce

$$2M_0^* + M_1^* = 4\pi(1 - \cos \rho). \tag{44}$$

We observe that if the circle C of radius ρ contains in its interior the curve $\mathfrak{L}$, it contains also the arc D. Hence $M_0 \leq M_0^*$. Likewise if C cuts $\mathfrak{L}$ in only one point, it has one of its extremities in the interior and the other in the exterior and so the arc D cuts the circle C also at only one point, that is to say, $M_1 \leq M_1^*$. It follows that, by (41) and (44),

$$\begin{aligned} 2F - 4L \sin \rho &\leq 2M_0^* + M_1^* - (M_3 + 2M_4 + \cdots) \\ &= 4\pi(1 - \cos \rho) - (M_3 + 2M_4 + \cdots); \end{aligned}$$

hence

$$F + \tfrac{1}{2}(M_3 + 2M_4 + \cdots) \leq 2\pi(1 - \cos \rho) + 2L \sin \rho. \tag{45}$$

This inequality implies (38).

The equality in (38) will be verified only if $M_i = 0$ for $i \geq 3$ and moreover $M_0 = M_0^*$, $M_1 = M_1^*$. The condition $M_i = 0$ for $i \geq 3$ carries with it $M_1 = M_1^*$; since in the case when the circle C cuts in only one point the arc of the great circle which joins the extremities of $\mathfrak{L}$, it must cut $\mathfrak{L}$ in an odd number of points. Consequently, the conditions for equality are:

(i) $M_i = 0$ (for $i \geq 3$). The curve $\mathfrak{L}$ cannot be cut by the circle C in more than two points.

(ii) $M_0 = M_0^*$, that is to say, if the circle C contains in its interior the two extremities of the curve $\mathfrak{L}$, it contains also the whole curve.

In particular, if the given curve $\mathfrak{L}$ is closed, the equality in (38) is valid only in the case of reduction to a point.

14. **Isoperimetric inequality on the sphere.** Let K be a convex curve on the sphere of unit radius. We consider the exterior parallel curve to K at the distance $\rho \leq \frac{1}{2}\pi$. This curve cannot have double points and its area is easy to calculate. The area is [3; 81]

$$S = F + L \sin \rho + 2\pi(1 - \cos \rho) - F(1 - \cos \rho), \tag{46}$$

or, with the values (36) of the area and the length of the circle of radius ρ,

$$S = F + F_0 + \frac{1}{2\pi}(LL_0 - FF_0). \tag{47}$$

Let us put, as in the last section, M_i $(i = 0, 2, 4, 6, \cdots)$ for the area covered by the centers of the circles of radius ρ which have i points in common with K (M_0 will be the area covered by the centers of the circles of radius ρ each of which contains K in its interior or which is contained in the interior of K). Since K is a closed curve, i is always even.

The expression (47) is equivalent to

$$M_0 + M_2 + M_4 + \cdots = F + F_0 + \frac{1}{2\pi}(LL_0 - FF_0) \tag{48}$$

and formula (37) gives

$$M_2 + 2M_4 + 3M_6 + \cdots = \frac{1}{\pi} LL_0 . \tag{49}$$

Let us consider a radius ρ such that $M_0 = 0$, that is, such that the circle of radius ρ neither can be totally interior to K nor can contain K in its interior. From (48) and (49) we deduce then

$$M_4 + 2M_6 + \cdots = \frac{1}{2\pi}(LL_0 + FF_0) - (F + F_0). \tag{50}$$

We observe that, by (36), $L_0^2 + F_0^2 - 4\pi F_0 = 0$; hence we can write the identity

$$\begin{aligned} &\frac{1}{2\pi}(LL_0 + FF_0) - (F + F_0) \\ &\qquad = \frac{1}{4\pi}[(L^2 + F^2 - 4\pi F) - (L - L_0)^2 - (F - F_0)^2] \end{aligned} \tag{51}$$

and (50) gives

$$L^2 + F^2 - 4\pi F = (L - L_0)^2 + (F - F_0)^2 + 4\pi(M_4 + 2M_6 + \cdots). \tag{52}$$

Since the second member of this equality always ≥ 0, we obtain the classical isoperimetric inequality on the sphere

$$L^2 + F^2 - 4\pi F \geq 0. \tag{53}$$

This inequality has often been proved. See [1], [3] and [2], and the bibliography in [4; 113]. For proof with methods of integral geometry analogous to those we follow in this paper, see [2; 83].

Equality (52) is valid when F_0 and L_0 are the area and length of any circle which neither contains K in its interior nor is contained in the interior of K. In particular, if C_0^* is the smallest circle which contains K in its interior and C_0 is the greatest circle which is contained in K, by neglecting the non-negative sum $M_4 + 2M_6 + \cdots$, we have

$$L^2 + F^2 - 4\pi F \geq (L - L_0)^2 + (F - F_0)^2, \tag{54}$$

$$L^2 + F^2 - 4\pi F \geq (L_0^* - L)^2 + (F_0^* - F)^2. \tag{55}$$

Taking into account the general inequality

$$x^2 + y^2 \geq \tfrac{1}{2}(x + y)^2, \tag{56}$$

we may combine inequalities (54) and (55) into the inequality

$$L^2 + F^2 - 4\pi F \geq \left(\frac{L_0^* - L_0}{2}\right)^2 + \left(\frac{F_0^* - F_0}{2}\right)^2. \tag{57}$$

This is a better form than (53) for the isoperimetric inequality.

If we substitute for L_0, L_0^*, F_0, F_0^* their values (36), relation (57) gives

$$L^2 + F^2 - 4\pi F \geq 4\pi^2 \sin^2 \frac{r_M - r_m}{2}, \tag{58}$$

where r_M and r_m are the spherical radii of the circles C_0^* and C_0.

T. Bonnesen [3; 82] has obtained the inequality

$$L^2 + F^2 - 4\pi F \geq 4\pi^2 \tan^2 \frac{r_M - r_m}{2},$$

which is better than our (58). His proof is completely different from ours.

For a sphere of radius R, inequality (57) takes the form

$$L^2 - 4\pi F + \left(\frac{F}{R}\right)^2 \geq \left(\frac{L_0^* - L_0}{2}\right)^2 + \left(\frac{F_0^* - F_0}{2R}\right)^2, \tag{59}$$

which as $R \to \infty$ gives the inequality

$$L^2 - 4\pi F \geq \left(\frac{L_0^* - L_0}{2}\right)^2 = \pi^2(r_M - r_m)^2, \tag{60}$$

which is a well-known isoperimetric inequality for plane curves established by Bonnesen [3; 63], [4; 113].

15. **An upper limitation for the isoperimetric deficit of convex spherical curves.** We now consider only convex curves K with *continuous radius of spherical curvature*. We understand by radius of spherical curvature the limit of the spherical radius of the circle which has three points in common with the curve as these points approach coincidence. This radius ρ ($\rho \leq \frac{1}{2}\pi$) is connected with the radius of geodesic curvature ρ_g by

$$\rho_g = \tan \rho.$$

Let ρ_M be the greatest radius and ρ_m the smallest radius of spherical curvature (both $\leq \frac{1}{2}\pi$). We wish to prove that

$$L^2 + F^2 - 4\pi F \leq \left(\frac{L_0^* - L_0}{2} + \frac{F_0^* - F_0}{2}\right)^2, \tag{61}$$

where L_0, F_0, L_0^*, F_0^* are now the lengths and areas of the circles whose radii are ρ_m and ρ_M respectively.

Likewise, as the area of the *exterior* parallel curve to K at distance ρ was expressed by (46), when we consider the *interior* parallel curve to K at a distance $\rho \leq \rho_m$, this curve will not have double points and its area is equal to

$$- L \sin \rho + F \cos \rho + 2\pi(1 - \cos \rho). \tag{62}$$

If we take $\rho = \rho_m$, area (62) will be the area covered by the centers of the circles of radius ρ_m which are contained in the interior of the convex curve K. If we represent this area by M_0, we can write

$$M_0 = - L \sin \rho_m + F \cos \rho_m + 2\pi(1 - \cos \rho_m). \tag{63}$$

We now wish to find the value of the area covered by the centers of the circles of radius ρ_M each of which contains K entirely in its interior. For this purpose we note that when the circle of radius ρ_M contains K in its interior, by a "dual" transformation (§2) the transformed circle (of radius $\frac{1}{2}\pi - \rho_M$) will be contained in the interior of the transformed curve K^* (whose length and area are $2\pi - F$ and $2\pi - L$ respectively). The area covered by the centers of the circles of radius ρ_M each of which contains K in its interior will then be given by (62) if we substitute ρ for $\frac{1}{2}\pi - \rho_M$, F for $2\pi - L$, and L for $2\pi - F$.

It follows that this area is given by

$$M_0^* = -L \sin \rho_M + F \cos \rho_M + 2\pi(1 - \cos \rho_M). \tag{64}$$

This has the same form as (63).

Let L_0, F_0 and L_0^*, F_0^* be the lengths and areas of the circles of radius ρ_m and ρ_M respectively, given by (36). Formulas (63) and (64) take the form

$$M_0 = F + F_0 - \frac{1}{2\pi}(LL_0 + FF_0) \tag{65}$$

and

$$M_0^* = F + F_0^* - \frac{1}{2\pi}(LL_0^* + FF_0^*). \tag{66}$$

When we take into account identity (51), these equalities give

$$L^2 + F^2 - 4\pi F = (L - L_0)^2 + (F - F_0)^2 - 4\pi M_0, \tag{67}$$

$$L^2 + F^2 - 4\pi F = (L_0^* - L)^2 + (F_0^* - F)^2 - 4\pi M_0^*. \tag{68}$$

Since M_0 and M_0^* are non-negative, we have

$$L^2 + F^2 - 4\pi F \leq (L - L_0)^2 + (F - F_0)^2, \tag{69}$$

$$L^2 + F^2 - 4\pi F \leq (L_0^* - L)^2 + (F_0^* - F)^2. \tag{70}$$

These inequalities give a first upper limit for the isoperimetric deficit $L^2 + F^2 - 4\pi F$.

From inequalities (69) and (70) we find

$$L^2 + F^2 - 4\pi F \leq (L - L_0 + F - F_0)^2. \tag{71}$$

$$L^2 + F^2 - 4\pi F \leq (L_0^* - L + F_0^* - F)^2. \tag{72}$$

Since the left sides are non-negative by (53) and since

$$xy \leq \left(\frac{x + y}{2}\right)^2,$$

by multiplication of (71) and (72), we find

$$L^2 + F^2 - 4\pi F \leq \left(\frac{L_0^* - L_0}{2} + \frac{F_0^* - F_0}{2}\right)^2. \tag{73}$$

For a sphere of radius R we have

$$L^2 - 4\pi F + \frac{F^2}{R^2} \leq \left(\frac{L_0^* - L_0}{2} + \frac{F_0^* - F_0}{2R}\right)^2, \tag{74}$$

and as $R \to \infty$,

$$L^2 - 4\pi F \leq \tfrac{1}{4}(L_0^* - L_0)^2 = \pi^2(\rho_M - \rho_m)^2, \tag{75}$$

where ρ_M and ρ_m are the greatest and the smallest radii of curvature of the plane convex curve K of length L and area F.

This inequality (75) is a known inequality obtained by Bottema [5]; see also [4; 83].

Bibliography

1. F. Bernstein, *Über die isoperimetrische Eigenschaft des Kreises auf der Kugeloberfläche und in der Ebene*, Mathematische Annalen, vol. 60(1905), pp. 117–136.
2. W. Blaschke, *Vorlesungen über Integralgeometrie*, Hamburger Mathematische Einzelschriften, Leipzig und Berlin, 1935.
3. T. Bonnesen, *Les problèmes des isopérimètres et des isépiphanes*, Gauthier-Villars, Paris, 1929.
4. T. Bonnesen and W. Fenchel, *Theorie der konvexen Körper*, Ergebnisse der Mathematik, Berlin, 1934, part I.
5. O. Bottema, *Eine obere Grenze für das isoperimetrische Defizit einer ebenen Kurve*, Proceedings of the Köninklijke Akademie van Wetenschappen, vol. 36(1933), pp. 442–446.
6. W. Crofton, *On the theory of local probability*, Philosophical Transactions of the Royal Society of London, vol. 158(1868), pp. 181–199.
7. W. Crofton, *Probability*, Encyclopaedia Britannica, 9th edition, 1885.
8. E. Czuber, *Geometrische Wahrscheinlichkeiten und Mittelwerte*, Leipzig, 1884.
9. R. Deltheil, *Probabilités géométriques* (*Traité du Calcul des Probabilités et de ses applications*, published under the direction of E. Borel), vol. II, fasc. II, Gauthier-Villars, Paris, 1926, pp. 74–76.
10. *Encyclopédie des Sciences Mathématiques*, vol. IV, 1904, pp. 26–27.
11. H. Hornich, *Eine allgemeine Ungleichung für Kurven*, Monatshefte für Mathematik und Physik, vol. 47(1939), pp. 432–438.
12. H. Lebesgue, *Exposition d'un mémoire de M. W. Crofton*, Nouvelles Annales de Mathématiques, (4), vol. 12(1912), pp. 481–502.
13. L. A. Santaló, *A theorem and an inequality referring to rectifiable curves*, American Journal of Mathematics, vol. 63(1941), pp. 635–644.

Instituto de Matemáticas, Rosario, Argentine Republic.

RELATED GENERA OF QUADRATIC FORMS

By Burton W. Jones

1. **Introduction.** The purpose of this paper is to provide formulas connecting the generalized representation function of a number by one genus of forms with the corresponding function of a divisor of this number by a related genus of forms. This result may be used as a recursion formula to simplify the evaluation of this function.

The generalized representation function was first developed by H. J. S. Smith and Minkowski and has lately been generalized, simplified and made more accessible by C. L. Siegel [17], [18]. Before describing the results of this paper, a few notions must be clarified. In the first place, it is understood throughout this paper that *every quadratic form* with which we deal has a non-vanishing determinant and that its matrix has rational integral elements.

We recall the following definitions: Two quadratic forms are said to be *equivalent*, that is, of the same *class*, if there is a unimodular transformation taking one into the other. We use "unimodular" in Siegel's sense to denote a linear transformation with integral coefficients and determinant ± 1. The classical definition of class differs from ours in that the determinant of the transformation is there required to be $+1$. Following Siegel we call such a transformation "properly unimodular". For example, we say that the forms $\alpha_1 = 4x^2 + 2xy + 5y^2$ and $\alpha_2 = 4x^2 - 2xy + 5y^2$ are equivalent but not properly equivalent and hence are of the same class but not of the same *proper class*. Similarly, in listing the automorphs (or units) we follow Siegel in including those of determinant -1 as well as those of determinant $+1$. This difference is of importance only with forms of an even number of variables since minus the identity transformation is always an automorph of determinant -1 when the number of variables is odd; hence, in such cases, our number of automorphs is twice the classical number and the number of classes is the same.

We say that two quadratic forms are in the same *genus* if they have the same signature (or index) and if, for every integer q, each form may be taken into the other by a linear transformation whose elements are rational numbers of denominators prime to q. This definition is equivalent to the original one [20], [14] phrased in terms of quadratic characters of numbers represented by a form and its concomitants. The equivalence of the two definitions was proved by H. J. S. Smith [20; 480ff] for ternary forms and stated by him [20; 516] for forms in n variables. Minkowski [14; 221] stated that his methods could be used to give a proof. The first complete published proof was given recently by C. L.

Received January 12, 1942; in revised form September 11, 1942. The basis for this paper was laid during a Sabbatic leave from Cornell University spent as a member of the Institute for Advanced Study. The author here wishes to express his thanks to both institutions.

Siegel [19]. (Strictly speaking, Siegel proved that two forms are of the same genus in the sense of the definition we here use if and only if they satisfy the conditions of Lemma 4 in this paper; Minkowski proved that the original definition is equivalent to the conditions of Lemma 4.) The determinant of any linear transformation with rational coefficients taking a form into any other of the same genus must have ± 1 as its determinant since the above definition implies that the determinants of the two forms are equal.

A quadratic form is said to be *properly primitive* if 1 is the g.c.d. of its coefficients and *improperly primitive* if it is not properly primitive and 1 is the g.c.d. of the elements of its matrix. A *primitive column matrix* or vector is one whose components have 1 as their g.c.d.

Siegel's principal result for positive forms is embodied in the following formula for quadratic forms of m and n variables, respectively, with matrices $\mathfrak{S}$ and $\mathfrak{T}$:

$$A_0(\mathfrak{S}, \mathfrak{T}) = \lambda \prod_p A_q(\mathfrak{S}, \mathfrak{T}) q^{\frac{1}{2}n(n+1)-mn},$$

where $A_0(\mathfrak{S}, \mathfrak{T}) = M(\mathfrak{S}, \mathfrak{T})/M(\mathfrak{S})$, $M(\mathfrak{S}, \mathfrak{T}) = \sum A(\mathfrak{S}_k, \mathfrak{T})/E(\mathfrak{S}_k)$, and $M(\mathfrak{S}) = \sum 1/E(\mathfrak{S}_k)$, the sums being over all classes of the genus, $E(\mathfrak{S}_k)$ being the number of integral automorphs of $\mathfrak{S}_k$, and $A(\mathfrak{S}_k, \mathfrak{T})$ being the number of solutions of $\mathfrak{X}'\mathfrak{S}\mathfrak{X} = \mathfrak{T}$ for $\mathfrak{X}$ a matrix with integral elements. $A_0(\mathfrak{S}, \mathfrak{T})$ is thus a kind of weighted mean of the number of representations of $\mathfrak{T}$ by the forms of the genus of $\mathfrak{S}$. On the right of the equality $A_q(\mathfrak{S}, \mathfrak{T})$ is the number of solutions (mod q) of $\mathfrak{X}'\mathfrak{S}\mathfrak{X} \equiv \mathfrak{T}$ (mod q), q is a sufficiently high power of p, and λ is a determined constant depending only on m, n and the determinants of $\mathfrak{S}$ and $\mathfrak{T}$. The product is over all primes p. Notice that $A_0(\mathfrak{S}, \mathfrak{T}) = A(\mathfrak{S}, \mathfrak{T})$ when there is but one form in the genus.

In a second paper [18], Siegel finds the corresponding result for indefinite forms. Since there is an infinite number of automorphs of an indefinite form, $M(\mathfrak{S}, \mathfrak{T})$ is replaced by a sum of certain functions over all solutions of $\mathfrak{X}'\mathfrak{S}\mathfrak{X} = \mathfrak{T}$ in which no one can be obtained from any other by multiplication on the left by an automorph. The right side of the formula is the same.

The "sufficiently high power of p" referred to is $p^a = q$, where $a > 2b$ and p^b is the highest power of p dividing $2T$, T being the determinant of $\mathfrak{T}$. The infinite product over all primes p not dividing $2ST$ (S being the determinant of $\mathfrak{S}$) has a simple finite expression. If $n = 1$, the same can be said for all odd primes not common factors of S and T. Hence, except for a finite but usually laborious computation, there is an explicit expression for the product on the right.

In this paper we deal solely with the case $n = 1$ when Siegel's expression reduces to formula (1) in §4 of this paper. The weighted mean $A_0(\mathfrak{S}, t)$ seems to be the closest one can, in general, come to an expression for the number of representations of a number by a form in a genus of more than one class. (The restriction $n = 1$ may not be entirely necessary but the most interesting applications of the theory of this paper are for this case.) If p is a prime factor of t, we establish for positive forms a simple relationship between $A_0(\mathfrak{S}, t)$ and the

functions $A_0(\mathfrak{W}, t/p)$ and $A_0(\mathfrak{D}, t/p)$, where ω and δ, of matrices $\mathfrak{W}$ and $\mathfrak{D}$ respectively, are representatives of so-called "related genera" of the form σ of matrix $\mathfrak{S}$; except, when $p = 2$, $A_0(\mathfrak{W}, t/p)$ is replaced by the sum of three such functions. As might be expected, the case $p = 2$ causes by far the most trouble. If p is a factor of S, we can, by continuing this process, arrive at an expression for $A_0(\mathfrak{S}, t)$ in terms of similar functions of other forms in which t is prime to p, thereby eventually avoiding some of the laborious computation referred to in the previous paragraph. If p is not a factor of S, $\mathfrak{D} = p\mathfrak{S}$ and we have a direct relationship between representations of tp and, if $t \equiv 0 \pmod p$, of t/p by the genus of σ, and of t by the genus of ω. These results are embodied in formulas (11) and (12). The details of this process may be seen in the examples at the end of the paper, especially Example 2 for ternary quadratic forms. The most difficult part of this paper consists in showing that each step of the process yields, except when $p = 2$, only forms of two genera represented by δ and ω. Again, the untidy part of the proof is for $p = 2$. Since the relationship referred to comes by way of $A_q(\mathfrak{S}, t)$, there will be exactly the *same relationship for the expression for indefinite forms* corresponding to $A_0(\mathfrak{S}, t)$ for positive forms.

An upper bound for the number of classes in the related genera is given in Theorems 7 and 8. This bound is, in certain cases, actually attained. This result is of considerable use in determining the number of classes in genera of forms.

By way of illustration we also exhibit explicit expressions for the value of the generalized representation function for positive ternary and quaternary forms. As far as the author knows, they have not yet appeared in print, though they are derivable for numbers prime to the determinant from results of H. J. S. Smith [20; 491].

The gist of this paper may be understood by omitting the case $p = 2$ throughout, reading §§1, 2 and 3 through the proof of Theorem 1, and reading Theorem 6 and the few lines immediately preceding it, Theorems 7 and 8, formulas (1), (11), (22), (29), (30) and such of the examples as prove interesting and enlightening.

2. **Notations and preliminary lemmas.** We adopt Siegel's notations in using German capitals for matrices and corresponding italic capitals for their determinants. We use corresponding lower case Greek letters for the quadratic forms. Lower case italic letters are rational integers unless otherwise specified. German lower case letters are column vectors which are called *integral* if their elements are integral. $A(\mathfrak{S}, \mathfrak{T})$, $A_q(\mathfrak{S}, \mathfrak{T})$ and $E(\mathfrak{S})$ are defined in the introduction. $\mathfrak{A} = \mathfrak{A}^{(m,n)}$ means that $\mathfrak{A}$ has m rows and n columns. If $\mathfrak{A} = \mathfrak{A}^{(m,n)}$, $m > n$ and $\mathfrak{C} = \mathfrak{C}^{(m,m-n)}$, then the matrix $(\mathfrak{A}\ \mathfrak{C})$ is that one whose first n columns are those of $\mathfrak{A}$ and whose last $m - n$ columns those of $\mathfrak{C}$. One similarly defines

$$\begin{pmatrix} \mathfrak{A}' \\ \mathfrak{C}' \end{pmatrix},$$

the primes denoting "transpose". $\mathfrak{N}$ is reserved to stand for a null matrix, $\mathfrak{n}$ for a null vector and $\mathfrak{E}$ for the identity matrix. If two forms σ and τ are equivalent, we write $\sigma \sim \tau$; if they are of the same genus, $\sigma \vee \tau$.

In order to make his paper self-containing, Siegel proved some lemmas and theorems which had previously been proved by Smith and Minkowski. Since the reader may find Siegel's proofs more readable, the page reference to his article as well as to the original proofs are given. We shall rephrase or specialize some of the lemmas in order to exhibit the exact form which we shall use.

LEMMA 1. *If a matrix* $\mathfrak{A} = \mathfrak{A}^{(m,n)}$ *with* $n < m$ *has integral elements and if* 1 *is the* g.c.d. *of its n-rowed minors, then there is a matrix* $\mathfrak{C} = \mathfrak{C}^{(m,m-n)}$ *with integral elements such that* $(\mathfrak{A}\ \mathfrak{C})$ *is unimodular. In fact,* $(\mathfrak{A}\ \mathfrak{C})$ *may be made properly unimodular* [17; 533], [20; 367–409].

LEMMA 2. *If q is an integer and* $\mathfrak{B}$ *a matrix of determinant* ± 1 *whose elements are rational numbers with denominators prime to q, then there is a unimodular matrix* $\mathfrak{B}_1$ *such that* $\mathfrak{B}_1 \equiv \mathfrak{B} \pmod{q}$ [17; 534].

LEMMA 3. *If q is an integer and* $\mathfrak{U}$ *a matrix of determinant* $\equiv \pm 1 \pmod{q}$ *and having integral elements, then there is a unimodular matrix* $\mathfrak{B} \equiv \mathfrak{U} \pmod{q}$ [17; Lemma 6], [20; vol. II, 635].

We use the following well-known

LEMMA 4. *For any power of an odd prime and hence for any odd number q, and any form* σ, *there exists a form* $\tau \sim \sigma$ *such that* $\mathfrak{T}$ *is a diagonal matrix* mod q. (See [17; 535]. In the statement of Siegel's lemma, R_p should be replaced by G_p.)

LEMMA 5. *Two forms* σ *and* τ *are of the same genus if and only if* $S = T$, *their indices are equal, and there is a form* $\tau_1 \sim \tau$ *such that* $\tau_1 \equiv \sigma \pmod{(2S)^3}$ [17; Lemma 20], [14; 71-79].

That Lemma 4 fails to hold for powers of 2 can be seen from the following example. $\sigma = 2x^2 + 2xy + 2y^2$ is of odd determinant and hence if it were of the same genus as a form $\equiv ax^2 + by^2 \pmod{2^m}$, a and b would need to be odd. This cannot be since σ represents no odd numbers and two equivalent forms represent the same numbers. As a matter of fact, it is true, though we shall not here prove it since we do not use it, that a form σ is equivalent to a diagonal form modulo an arbitrary power of 2 if and only if σ and all its concomitants are properly primitive. However, we do need lemmas which, for powers of 2, take the place of Lemma 4. Parts of Lemmas 6 have been proved by Minkowski and C. Jordan. (See [14; 16-21]. At least the basis for these results was earlier laid by Jordan [8]. A brief derivation of 6a and 6b is given in [16].)

LEMMA 6a. *If* α *is a properly primitive form and* $q \neq 1$, *an arbitrary but fixed power of* 2, *there exists a form* $\beta \sim \alpha$ *such that*

$$\beta \equiv 2^{r_1}\beta_1 + 2^{r_2}\beta_2 + \cdots + 2^{r_t}\beta_t \pmod{q},$$

where β_i are forms of odd determinant and no two have any variables in common. The exponents e_i are non-negative integers.

LEMMA 6b. *Every form of odd determinant is equivalent to a form of type α or type β below according as it is properly or improperly primitive:*

$$\alpha \equiv a_1x_1^2 + a_2x_2^2 + \cdots + a_sx_s^2 \pmod{q},$$

$$\beta \equiv \beta_1 + \beta_2 + \cdots + \beta_r \pmod{q},$$

where the a_i are odd integers and β_i are improperly primitive binary forms no two of which have a variable in common.

LEMMA 6c. *Every form $2ax^2 + 2bxy + 2cy^2$ in which b is an odd integer can be carried by an integral transformation of odd determinant into a form*

$$\beta \equiv 2xy \pmod{q} \quad \text{or } \beta \equiv 2x^2 + 2xy + 2y^2 \pmod{q}$$

according as ac is even or odd.

The proofs of Lemmas 6c and 6d are essentially those of Gordon Pall, being simpler than those which the author devised. In fact, Pall's contribution to the simplification of the whole case $p = 2$ is notable.

Proof. Since b is odd, we can solve $ax_1^2 + bx_1x_2 + cx_2^2 \equiv 1 \pmod{q}$. Taking x_1, x_2 as the first column of a linear transformation of odd determinant, we obtain an equivalent form of the type $x^2 + b_1xy + c_1y^2$, with b_1 odd. Replace x by $x + hy$, y by ky, and choose h and k so that

$$2h + b_1k \equiv 1, \qquad h_1^2 + b_1hk + c_1k^2 \equiv e \pmod{q},$$

where e is 0 or 1 according as c_1 is even or odd. The latter congruence is equivalent to $(2h + b_1k)^2 + (4c_1 - b_1^2)k^2 \equiv 4e \pmod{4q}$ and is solvable with k odd. If $e = 1$, our resulting form $\equiv x^2 + xy + y^2 \pmod{q}$, while if $e = 0$ it $= x^2 + xy$, which is equivalent $(\text{mod } q)$ to xy.

LEMMA 6d. *The form β of Lemma 6b is equivalent to a form*

$$\beta_0 \equiv 2(jx_1^2 + bx_1x_2 + jx_2^2) + 2x_3x_4 + \cdots + 2x_{2r-1}x_{2r} \pmod{q},$$

where j is 0 or 1 and b is odd.

Proof. The transformation $x_1 = y_1 + y_3 + y_4$, $x_2 = y_2 - y_3 - y_4$, $x_3 = -y_1 + y_2 - y_4$, $x_4 = -y_3 + y_4$ of determinant 3 replaces $x_1^2 + x_1x_2 + x_2^2 - x_3^2 - x_3x_4 - x_4^2$ by $3y_1y_2 + 3y_3y_4$. Hence Lemma 6c shows that a transformation of odd determinant takes β of Lemma 6b into a form β_0 with b an arbitrary odd integer. We can choose b to make the determinants of β and β_0 congruent $(\text{mod } q)$; then the accompanying transformation will have determinant ± 1 $(\text{mod } q)$. The use of Lemma 3 completes our proof.

3. **Related genera.** Let β be any quadratic form in r variables and let $\mathfrak{B}$ be its matrix. Let p be any prime and let the vector $\mathfrak{h}$ be a primitive solution of

$\mathfrak{x}'\mathfrak{B}\mathfrak{x} \equiv 0 \pmod{p}$. Write $\mathfrak{Z} = (p\mathfrak{Z}^*\ \mathfrak{h})$, where $\mathfrak{Z}^0 = (\mathfrak{Z}^*\ \mathfrak{h})$ is a unimodular matrix. It is easily seen that $\mathfrak{Z} = \mathfrak{Z}^0\mathfrak{P}$, where

$$\mathfrak{P} = \begin{pmatrix} p\mathfrak{E}^{(r-1)} & \mathfrak{n} \\ \mathfrak{n} & 1 \end{pmatrix}.$$

$\mathfrak{Z}$ takes β into a form $p\gamma$ of matrix $p\mathfrak{C}$, where

$$p\mathfrak{C} = \mathfrak{Z}'\mathfrak{B}\mathfrak{Z} = \begin{pmatrix} p^2\mathfrak{Z}^{*\prime}\mathfrak{B}\mathfrak{Z}^* & p\mathfrak{Z}^{*\prime}\mathfrak{B}\mathfrak{h} \\ p\mathfrak{h}'\mathfrak{B}\mathfrak{Z}^* & \mathfrak{h}'\mathfrak{B}\mathfrak{h} \end{pmatrix},$$

and C, the determinant of $\mathfrak{C}$, is $p^{r-2}B$.

Let $(\mathfrak{G}'\ \mathfrak{g})' = \mathfrak{Z}^{0-1}$. Then $(\mathfrak{G}'\ \mathfrak{g})'(\mathfrak{Z}^*\ \mathfrak{h}) = \mathfrak{E}^{(r)}$; whence $\mathfrak{G}\mathfrak{Z}^* = \mathfrak{E}^{(r-1)}$ and $\mathfrak{G}\mathfrak{h} = (\mathfrak{g}'\ \mathfrak{Z}^*)' = \mathfrak{n}$, $\mathfrak{g}'\mathfrak{h} = 1$. Also $\mathfrak{Z}^*\mathfrak{G} + \mathfrak{h}\mathfrak{g}' = \mathfrak{E}^{(r)}$.

Let $\mathfrak{x}' = (x_1, x_2, \cdots, x_r) \equiv s\mathfrak{h}' \pmod{p}$, where s is a scalar not divisible by p; and $\mathfrak{x} = \mathfrak{Z}\mathfrak{y}$. Then

$$\mathfrak{y} = \mathfrak{Z}^{-1}\mathfrak{x} = \mathfrak{P}^{-1}\mathfrak{Z}^{0-1}\mathfrak{x} = \mathfrak{P}^{-1}(\mathfrak{G}'\ \mathfrak{g})'\mathfrak{x} = \mathfrak{P}^{-1}\begin{pmatrix} \mathfrak{G}\mathfrak{x} \\ \mathfrak{g}'\mathfrak{x} \end{pmatrix}.$$

First, notice that $\mathfrak{g}'\mathfrak{x} \equiv \mathfrak{g}'s\mathfrak{h} \equiv s \pmod{p}$, which implies that $y_r \equiv s \pmod{p}$, where $\mathfrak{y} = (y_1, y_2, \cdots, y_r)'$. Second, $\mathfrak{G}\mathfrak{h} = \mathfrak{n}$ implies $\mathfrak{G}\mathfrak{x} \equiv \mathfrak{G}s\mathfrak{h} \equiv \mathfrak{n} \pmod{p}$ and hence $\mathfrak{y}$ is integral if $\mathfrak{x}$ is. Thus, if $\mathfrak{x}'\mathfrak{B}\mathfrak{x} = pa$, then $\mathfrak{y}'\mathfrak{Z}'\mathfrak{B}\mathfrak{Z}\mathfrak{y} = p\mathfrak{y}'\mathfrak{C}\mathfrak{y} = pa$ and $\mathfrak{y}'\mathfrak{C}\mathfrak{y} = a$. Since the process is reversible, we have established a 1-1 correspondence between each integral $\mathfrak{x}$ such that $\mathfrak{x}'\mathfrak{B}\mathfrak{x} = pa$ with $\mathfrak{x} \equiv s\mathfrak{h} \pmod{p}$ and each integral $\mathfrak{y}$ such that $\mathfrak{y}'\mathfrak{C}\mathfrak{y} = a$ with $y_r \equiv s \pmod{p}$.

We remark that for any r by r matrix $\mathfrak{M}$, $\mathfrak{P}^{-1}\mathfrak{M}\mathfrak{P}$ can be obtained from $\mathfrak{M}$ by multiplying its last row by p and its last column by $1/p$.

THEOREM 1. *If* $\mathfrak{h}_2 \equiv s\mathfrak{h}_1 \pmod{p}$, *where s is a scalar* $\not\equiv 0 \pmod{p}$, *and if the transformations* $\mathfrak{Z}$ *and the forms* γ *for* $\mathfrak{h}_1$ *and* $\mathfrak{h}_2$ *are labelled* $\mathfrak{Z}_1$, $\mathfrak{Z}_2$ *and* γ_1, γ_2, *respectively, then* $\gamma_1 \sim \gamma_2$.

Proof. $\mathfrak{Z}_1$ takes $\mathfrak{B}$ into $p\mathfrak{C}_1$ and $\mathfrak{Z}_2$ takes $\mathfrak{B}$ into $p\mathfrak{C}_2$. Hence $\mathfrak{Z}_1^{-1}\mathfrak{Z}_2$ takes $\mathfrak{C}_1$ into $\mathfrak{C}_2$. Now $\mathfrak{Z}_1^{-1}\mathfrak{Z}_2 = \mathfrak{P}^{-1}\mathfrak{Z}_1^{0-1}\mathfrak{Z}_2^0\mathfrak{P}$ and, by the above remark, to show $\mathfrak{Z}_1^{-1}\mathfrak{Z}_2$ an integral matrix, we need merely show that the first $r - 1$ numbers in the last column of $\mathfrak{Z}_1^{0-1}\mathfrak{Z}_2^0$ are divisible by p. Now

$$\mathfrak{Z}_1^{0-1}\mathfrak{Z}_2^0 = \begin{pmatrix} \mathfrak{G}_1 \\ \mathfrak{g}_1' \end{pmatrix}(\mathfrak{Z}_2^*\mathfrak{h}_2) = \begin{pmatrix} \mathfrak{G}_1\mathfrak{Z}_2^* & \mathfrak{G}_1\mathfrak{h}_2 \\ \mathfrak{g}_1'\mathfrak{Z}_2^* & \mathfrak{g}_1'\mathfrak{h}_2 \end{pmatrix},$$

$\mathfrak{G}_1$ and $\mathfrak{g}_1$ being defined above. We know that $\mathfrak{G}_1\mathfrak{h}_2 \equiv \mathfrak{G}_1\, s\mathfrak{h}_1 \equiv s\mathfrak{n} \pmod{p}$, and our proof is complete.

Our goal in this section is to prove that if $\mathfrak{h}_1$ and $\mathfrak{h}_2$ are two solutions of $\mathfrak{x}'\mathfrak{B}\mathfrak{x} \equiv 0 \pmod{p}$ then, except when $p = 2$, the corresponding forms γ_1 and γ_2 are of the same genus. In case $p = 2$ our results are not so simple but are no less definite. To this end we prove the following lemmas which the proofs of Theorems 2 and 3 require.

LEMMA 7. *Let $\mathfrak{B}$ be the matrix of a primitive form in r variables, p an odd prime not dividing B, and $\mathfrak{h}_1$ and $\mathfrak{h}_2$ two non-zero* (mod p) *vectors such that* $\mathfrak{h}_i'\mathfrak{B}\mathfrak{h}_i \equiv 0$ (mod p), $i = 1, 2$. *Then there exists a vector $\mathfrak{c}$ such that $\mathfrak{c}'\mathfrak{B}\mathfrak{c}$, $\mathfrak{c}'\mathfrak{B}\mathfrak{h}_1$ and $\mathfrak{c}'\mathfrak{B}\mathfrak{h}_2$ are all prime to p.*

Proof. First, notice that this lemma will hold for a form β if it holds for any form equivalent to β. Hence, by Lemma 4, we may take the matrix of β to be a diagonal matrix (mod q), where q is an arbitrary power of p. Then $\beta \equiv b_1x_1^2 + b_2x_2^2 + \cdots + b_rx_r^2$ (mod q), where all b_i are prime to p. Then, for each set of values of $x_2, \cdots, x_r$, there are at least $p - 2$ different values of x_1 such that $\mathfrak{x}'\mathfrak{B}\mathfrak{x} \not\equiv 0$ (mod p); hence not less than $(p-2)p^{r-1}$ different vectors $\mathfrak{c}$ such that $\mathfrak{c}'\mathfrak{B}\mathfrak{c} \not\equiv 0$ (mod p). On the other hand, for any $\mathfrak{B}\mathfrak{h} \not\equiv \mathfrak{n}$ (mod p), there are exactly p^{r-1} vectors $\mathfrak{c}$ such that $\mathfrak{c}'\mathfrak{B}\mathfrak{h} \equiv \mathfrak{n}$ (mod p) for if, for instance, the leading element of $\mathfrak{B}\mathfrak{h}$ is not divisible by p, there will be, for every $c_2, \cdots, c_r$, exactly one component c_1 of $\mathfrak{c}$ satisfying the congruence. Now $\mathfrak{B}\mathfrak{h}_i \not\equiv \mathfrak{n}$ (mod p) for $i = 1$ and 2 since $\mathfrak{h}_i$ is non-zero and B is prime to p. Hence, there is some $\mathfrak{c}$ satisfying the conditions of the lemma if

$$(p-2)p^{r-1} - 2p^{r-1} = p^{r-1}(p-4) > 0,$$

i.e., if $p \geq 5$.

It remains to consider $p = 3$. Then, by renumbering the variables if necessary, we can write

$$\beta \equiv x_1^2 + x_2^2 + \cdots + x_v^2 - x_{v+1}^2 - \cdots - x_r^2 \qquad \text{(mod 3)}.$$

Write $\mathfrak{h}_i' = (h_{i1}, h_{i2}, \cdots, h_{ir})$. First, if for some k, $h_{1k}\,h_{2k} \not\equiv 0$ (mod 3), take $\mathfrak{c}' = (0, \cdots, 0, 1, 0, \cdots, 0)$, where the 1 is the k-th component. Second, if for some k and $l \leq v$ or k and $l > v$ it is true that $h_{1k}h_{2l} \not\equiv 0$ (mod p) but $h_{1k}h_{2k} \equiv h_{1l}h_{2l} \equiv 0$ (mod p), take $\mathfrak{c}' = (0, \cdots, 0, 1, 0, \cdots, 0, 1, 0, \cdots, 0)$, where the two non-zero components are the k-th and l-th. Third, if $h_{1k} \equiv 0$ (mod p) for all $k > v$ and $h_{2k} \equiv 0$ (mod p) for all $k \leq v$, we notice that $\mathfrak{h}'\mathfrak{B}\mathfrak{h} \equiv 0$ (mod 3) implies that two h_{1l}, h_{1m}, for $m, l \leq v$, $\not\equiv 0$ (mod 3) and two h_{2w}, h_{2u}, with $w, u > v$, $\not\equiv 0$ (mod 3). Then take as $\mathfrak{c}$ the vector whose only non-zero components are 1's in the l, m, and w places. We have considered essentially all cases.

The reader who wishes to avoid the laborious case $p = 2$ may pass immediately to Theorems 2 and 3.

For $p = 2$, we use the following terminology; if $\mathfrak{h}$ is a vector (mod 2), we call $(1, 1, \cdots, 1)' - \mathfrak{h}$ its *complementary vector* or *complement* and denote it by $\mathfrak{h}^0$.

LEMMA 8. *If* $\beta \equiv x_1^2 + \cdots + x_v^2 - x_{v+1}^2 - \cdots - x_r^2$ (mod 4) *and $\mathfrak{h}_1$ and $\mathfrak{h}_2$ are two primitive vectors neither of which has all its components* 1, *then, with the exceptions below, there always exists a vector $\mathfrak{c}$ such that $\mathfrak{c}'\mathfrak{B}\mathfrak{h}_1$ and $\mathfrak{c}'\mathfrak{B}\mathfrak{h}_2$ are odd (that is, $\mathfrak{c}'\mathfrak{h}_1$ and $\mathfrak{c}'\mathfrak{h}_2$ are odd) while* $\mathfrak{c}'\mathfrak{B}\mathfrak{c} \equiv 2$ (mod 4). *In each vector the semicolon appears after the v-th component. Exceptional values of $\mathfrak{h}_1'$ and $\mathfrak{h}_2'$* (mod 2) *are*

1. (0; 1, 1) *and* (1; 1, 0); (0; 1, 1) *and* (0; 1, 1);
2. (0, 0; 1, 1) *and one of* (0, 0; 1, 1), (1, 1; 0, 0), (0, 1; 0, 1);
3. (0, 0; 1, 1, 0) *and one of* (1, 1; 0, 0, 0), (0, 1; 1, 1, 1);
4. (0, 0, 0; 1, 1, 0) *and* (1, 1, 0; 0, 0, 0);
5. (0, 0, 0; 1) *and* (0, 1, 1; 0); (0, 0; 1, 1) *and any* $\mathfrak{h}_2$ *not in* 2; (0, 0; 0, 1) *and* (0, 1; 0, 0); (0, 1) *and* (1, 0);
6. *vectors obtained from the above by some or all of the following changes: interchanging* $\mathfrak{h}_1'$ *and* $\mathfrak{h}_2'$, *permuting the variables of both* $\mathfrak{h}_1$ *and* $\mathfrak{h}_2$ *by the same permutation, replacing one* $\mathfrak{h}$ *or both by its complement, multiplying* β *by* -1.

Exceptions 5 contribute no cases in which $\mathfrak{h}_i'\mathfrak{h}_i$ *is even.*

Pall has contributed considerably to the simplication of the proof of this lemma and of Lemmas 9 and 10.

Proof. First, we show that if the lemma holds for $\mathfrak{h}_1$ it holds for $\mathfrak{h}_1^0$. $\mathfrak{c}'\mathfrak{B}\mathfrak{h}_1^0 = \mathfrak{c}'\mathfrak{B}(1, 1, \cdots, 1)' - \mathfrak{c}'\mathfrak{B}\mathfrak{h}_1$ while $\mathfrak{c}'\mathfrak{B}(1, 1, \cdots, 1)' \equiv \mathfrak{c}'(1, 1, \cdots, 1)' \equiv \mathfrak{c}'\mathfrak{c} \equiv \mathfrak{c}'\mathfrak{B}\mathfrak{c} \equiv 0 \pmod 2$. Note also that $\mathfrak{h}_1^{0\prime}\mathfrak{B}\mathfrak{h}_1^0 \equiv r + \mathfrak{h}_1'\mathfrak{B}\mathfrak{h}_1 \pmod 2$.

In what follows we use asterisks to indicate unknown components. For instance, $(0, \cdots, 0; 1, 1, *, \cdots, *)$ is a vector whose first v components are 0, whose $(v + 1)$-th and $(v + 2)$-th are 1 and whose others are unknown. We write such a symbol even if $v = 0$ or $r - v = 2$ or both though the exhibition of two 1's indicates $r - v \geq 2$.

Now consider what vectors $\mathfrak{h}_1$ and $\mathfrak{h}_2$ can be treated by the simplest vectors $\mathfrak{c}$ such that $\mathfrak{c}'\mathfrak{B}\mathfrak{c} \equiv 2 \pmod 4$. If $\mathfrak{c}$ has every component zero except two components c_j and c_k such that j and k both $\leq v$ or both $> v$, then $\mathfrak{c}'\mathfrak{h}_1 = h_{1j} + h_{1k}$ and $\mathfrak{c}'\mathfrak{h}_2 = h_{2j} + h_{2k}$. We call such a $\mathfrak{c}$ *of type* I. It will take care of all $\mathfrak{h}_1$ and $\mathfrak{h}_2$ in which not all of $h_{11}, \cdots, h_{1v}$ are equal and not all of $h_{21}, \cdots, h_{2v}$ are equal, and of all $\mathfrak{h}_1$ and $\mathfrak{h}_2$ in which not all of $h_{1,v+1}, \cdots, h_{1r}$ are equal and not all of $h_{2,v+1}, \cdots, h_{2r}$ are equal; for, any such pair will have a 0 and 1 corresponding to a 0 and 1, or to a 1 and 0, both on the same side of the semicolon. Using complements if necessary, we have left only two possibilities:

(a) $\mathfrak{h}_1' = (0, \cdots, 0; 1, \cdots, 1)$ and $\mathfrak{h}_2'$ arbitrary,

(b) $\mathfrak{h}_1' = (0, \cdots, 0; 1, 0, *, \cdots, *)$ and $\mathfrak{h}_2' = (*, \cdots, *, 0, 1; 0, \cdots, 0)$.

The next simplest vectors $\mathfrak{c}$ are symbolized by (1, 1, 1; 1) or by (1; 1, 1, 1) and will eliminate all pairs $\mathfrak{h}_1$ and $\mathfrak{h}_2$ for which

$$\text{(c)} \qquad h_{1h} + h_{1k} + h_{1l} + h_{1m} \equiv h_{2h} + h_{2k} + h_{2l} + h_{2m} \equiv 1 \pmod 2$$

for any choice of subscripts three on one side and one on the other side of the semicolon. If $v \geq 4$, the denial of (c) with $m = v + 1$ and $h, k, l \leq v$ implies $h_{2h} + h_{2k} + h_{2l} + h_{2,v+1} \equiv 0 \pmod 2$ for all choices of $h, k, l \leq v$ which, in turn, implies $h_{2h} \equiv h_{2k} \equiv h_{2l} \equiv h_{2,v+1} \pmod 2$. Then, replacing $\mathfrak{h}_2$ by its complement and permuting variables if necessary we can write $\mathfrak{h}_2' = (0, \cdots, 0; 1, *, \cdots, *)$,

which excludes (b) above and to which our "next simplest" $\mathfrak{c}$ applies. Since the cases $r - v \geq 4$ may be similarly handled, there remain the following cases:

$$\begin{array}{c}(0, 0, 0; 1, 1, 1)\\(*, *, *; *, *, *),\end{array}$$

where denial of (c) implies $\mathfrak{h}_2$, or its complement is 0,

$$\begin{array}{c}(0, 0, 0; 1, *, 0)\\(1, *, 0; 0, 0, 0),\end{array}$$

where denial of (c) implies that both asterisks are 1, yielding exception 4,

$$\begin{array}{c}(0, 0, 0; 1, 1)\\(a, b, c; d, e),\end{array}$$

where denial of (c) implies $a + b + c \equiv d \equiv e \pmod 2$ yielding exception 3.1,

$$\begin{array}{ccccc}(0, 0, 0; 1, 0) & (0, 0, 0; 1) & (0, 0; 1, 1) & (0, 0; 0, 1) & (0, 0; 1)\\(*, 0, 1; 0, 0), & (*, *, 1; 0), & (*, *; *, *), & (0, 1; 0, 0), & (*, 1; 0),\end{array}$$

and the pair (0; 1) and (1; 0) which are all easily dealt with.

LEMMA 9. *If* $\beta \equiv 2x_1 x_2 + 2x_3 x_4 + \cdots + 2x_{2s-1} x_{2s} \pmod 4$ $(s > 0)$ *and* $\mathfrak{h}_1$ *and* $\mathfrak{h}_2$ *are two non-zero* (mod 2) *vectors with* $2s$ *components each, there exists a vector* $\mathfrak{c}$ *such that* $\mathfrak{c}'\mathfrak{B}\mathfrak{c} \equiv 2 \pmod 4$ *and* $\mathfrak{c}'\mathfrak{B}\mathfrak{h}_1$ *and* $\mathfrak{c}'\mathfrak{B}\mathfrak{h}_2$ *are odd, with the following exceptions, where the* $\mathfrak{h}_i$ *are listed* (mod 2):

1. $s = 1$ *and either* $\mathfrak{h}_1' = (1, 0)$, $\mathfrak{h}_2' = (1, 1)$ *or* $\mathfrak{h}_1' = (1, 1) = \mathfrak{h}_2'$,
2. $s = 2$ *and* $\mathfrak{h}_1'$ *and* $\mathfrak{h}_2'$ *are* $(1, 0, 1, 1)$ *and* $(1, 1, 1, 0)$ *or* $(0, 0, 1, 1)$ *and* $(1, 1, 0, 0)$ *or* $(1, 0, 1, 1)$ *and* $(0, 0, 1, 1)$.

Proof. First, let $s = 1$. The vector $\mathfrak{c}' = (1, 1)$ disposes of $\mathfrak{h}_1$ and $\mathfrak{h}_2$ if each has the components 0, 1 in some order. Such a $\mathfrak{c}$ we say is of type I.

Second, if $s = 2$, write $\mathfrak{h}_1' = (a_1, a_2, a_3, a_4)$ and $\mathfrak{h}_2' = (b_1, b_2, b_3, b_4)$. A permutation of $\mathfrak{c}' = (1, 1, 1, 0)$, which we call of type II, disposes of $\mathfrak{h}_1$ and $\mathfrak{h}_2$ unless for every $a_i + a_j + a_k$ which is odd the corresponding $b_i + b_j + b_k$ is even. Hence, if $\mathfrak{h}_1' = (0, 0, 0, 1)$, every $\mathfrak{h}_2$ is eliminated except those for which $b_1 + b_2 + b_4 = b_1 + b_3 + b_4 = b_2 + b_3 + b_4 \equiv 0 \pmod 2$. Hence, $\mathfrak{h}_2' = (0, 0, 0, 0)$ or $(1, 1, 1, 0)$, the latter of which is eliminated by a $\mathfrak{c}$ of type I. Similarly, if $\mathfrak{h}_2$ is not the zero vector, all but the following typical cases are eliminated:

$$\begin{array}{l}\mathfrak{h}_1' = (0, 0, 1, 1) \text{ with } \mathfrak{h}_2' = (1, 1, 0, 0), (0, 1, 1, 1) \text{ or } (1, 0, 1, 1),\\ \mathfrak{h}_1' = (0, 1, 1, 1) \text{ with } \mathfrak{h}_2' = (1, 1, 0, 1),\end{array}$$

which are essentially listed in exception 2.

Third, if $s = 3$, we see that if the first four components of $\mathfrak{h}_1$ are neither of

those displayed above, our previous treatment eliminates all cases except when the first four components of $\mathfrak{h}_2$ are 0; then we interchange $\mathfrak{h}_1$ and $\mathfrak{h}_2$ and have to consider for $\mathfrak{h}_1'$

$$(0, 0, 1, 1, *, *), (0, 1, 1, 1, *, *), \text{ and } (0, 0, 0, 0, *, *).$$

The last permutes into one of the first two unless the last two components are 1, 0 in some order when the above paragraph disposes of all non-zero $\mathfrak{h}_2$. Similarly, a $\mathfrak{c}$ of type II disposes of all but the following:

$$\begin{matrix} (0, 0, 1, 1, 1, 1) & (0, 0, 1, 1, 0, 0) & (0, 0, 1, 1, 0, 0) \\ (1, 1, 0, 0, 0, 0), & (1, 1, 0, 0, 1, 1), & (0, 1, 1, 1, 1, 0), \end{matrix}$$

and $\mathfrak{c}' = (1, 0, 0, 1, 1, 1)$ disposes of these.

LEMMA 10. *Let β be a quadratic form of odd determinant of the type of Lemma* 8 *or β_0 of Lemma* 6d *and $\mathfrak{h}_1$ and $\mathfrak{h}_2$ two distinct non-zero solutions* (mod 2) *of* $\mathfrak{x}'\mathfrak{B}\mathfrak{x} \equiv 0$ (mod 2) *such that* $\mathfrak{h}_1'\mathfrak{B}\mathfrak{h}_1 \equiv \mathfrak{h}_2'\mathfrak{B}\mathfrak{h}_2$ (mod 4), *then, with the exceptions below, there exists a vector $\mathfrak{c}$ such that* $\mathfrak{c}'\mathfrak{B}\mathfrak{c} \equiv 2$ (mod 4) *while $\mathfrak{c}'\mathfrak{B}\mathfrak{h}_1$ and $\mathfrak{c}'\mathfrak{B}\mathfrak{h}_2$ are odd.*

Exceptions.

1. *β is of the form in Lemma* 8 *and $\mathfrak{h}_1$ and $\mathfrak{h}_2$ are*

$$\begin{matrix} (1, 1; 0, 0, 0) \\ (0, 0; 1, 1, 0) \end{matrix} \text{ or } \begin{matrix} (0, 0; 1, 1) \\ (1, 1; 0, 0) \end{matrix} \text{ or } \begin{matrix} (1, 1, 1; 0, 1) \\ (1, 1, 0; 0, 0) \end{matrix} \text{ or } \begin{matrix} (0, 0, 0; 0, 1, 1) \\ (1, 1, 0; 0, 0, 0) \end{matrix}$$

or pairs obtained from them by one or more of the following changes: permuting variables, replacing β by $-\beta$, replacing one or both of the last displayed pair by its complement.

2. *β is of the form in Lemma* 9 *and $\mathfrak{h}_1$ and $\mathfrak{h}_2$ are obtained by permuting the variables of the following pairs:*

$$\begin{matrix} (0, 1, 1, 1) & (0, 0, 1, 1) & (1, 0, 1, 1) \\ (1, 1, 1, 0), & (1, 1, 0, 0), & (0, 0, 1, 1). \end{matrix}$$

3. *β is of the form in Lemma* 8 *and one of $\mathfrak{h}_1$ and $\mathfrak{h}_2$ has all its components* 1.

Proof. When β is of the form of Lemma 8, the exceptions there contribute exceptions 1 and 3 of this lemma. When β is of the form of Lemma 9, our present exception 2 results. It remains to consider

$$\beta \equiv 2x_1^2 + 2x_1x_2 + 2x_2^2 + 2x_3x_4 + \cdots + 2x_{2r-1}2_{2r} \qquad (\text{mod } q),$$

where, as usual, q is an arbitrary power of 2. If the first two components of $\mathfrak{h}_1$ and of $\mathfrak{h}_2$ are different, we choose $\mathfrak{c} = (1, 1, 0, \cdots, 0)$ while if $\mathfrak{h}_1$ and $\mathfrak{h}_2$ each has 1 as its first component, $\mathfrak{c} = (1, 0, \cdots, 0)$ suffices. Hence, we need to consider only the typical case when the first two components of $\mathfrak{h}_1$ are 0. Now, Lemma 9 shows that $r = 2$ or 3 and we need to consider only the cases in which the last two or four components of $\mathfrak{h}_1$ and $\mathfrak{h}_2$ occur among the exceptions of that lemma.

First, if $r = 2$, $\mathfrak{h}_1'\mathfrak{B}\mathfrak{h}_1 \equiv \mathfrak{h}_2'\mathfrak{B}\mathfrak{h}_2 \pmod 4$ implies that either $\mathfrak{h}_1' = (0, 0, 1, 1) = \mathfrak{h}_2'$ or $\mathfrak{h}_1' = (0, 0, 1, 0)$ and $\mathfrak{h}_2' = (1, *, 1, 1)$ and in the latter case we take $\mathfrak{c}' = (0, 1, 1, 0)$.

Second, if $r = 3$, $\mathfrak{h}_1'\mathfrak{B}\mathfrak{h}_1 \equiv \mathfrak{h}_2'\mathfrak{B}\mathfrak{h}_2' \pmod 4$ implies that the first two components of $\mathfrak{h}_2$ are 0, 0 and the last four components of $\mathfrak{h}_1$ and $\mathfrak{h}_2$ are one of the pairs of exceptions 2 of Lemma 9. Thus, there remain:

$$\begin{matrix} (0, 0, 1, 0, 1, 1) & (0, 0, 0, 0, 1, 1) & (0, 0, 1, 0, 1, 1) \\ (0, 0, 1, 1, 1, 0), & (0, 0, 1, 1, 0, 0), & (0, 0, 0, 0, 1, 1), \end{matrix}$$

and $\mathfrak{c} = (1, 0, 0, *, 0, 1)$ suffices, where the asterisk is 1 in the second case and otherwise 0.

THEOREM 2. *Let β be a primitive (not necessarily properly primitive) form in n variables, p a prime not dividing B, $\mathfrak{h}_1$ and $\mathfrak{h}_2$ two primitive solutions of $\mathfrak{x}'\mathfrak{B}\mathfrak{x} \equiv 0$* (mod p) *with the further proviso that if $p = 2$, $\mathfrak{h}_1'\mathfrak{B}\mathfrak{h}_1 \equiv \mathfrak{h}_2'\mathfrak{B}\mathfrak{h}_2$* (mod 4). *Then, with the exception below, for q an arbitrary power of p, there exists a unimodular transformation $\mathfrak{A}$ such that $\mathfrak{A}'\mathfrak{B}\mathfrak{A} \equiv \mathfrak{B}$* (mod q) *and $\mathfrak{A}\mathfrak{h}_1 \equiv \mathfrak{h}_2$* (mod p).

The exception for $p = 2$ is: β properly primitive, n even, and either $\mathfrak{h}_1$ or $\mathfrak{h}_2$ having its components congruent in order to the elements in the principal diagonal of $\mathfrak{B}^{-1}$.

Proof. First, if $\mathfrak{h}_2'\mathfrak{B}\mathfrak{h}_1 \not\equiv 0 \pmod p$, let $\mathfrak{h}_3 = \mathfrak{h}_2 - \mathfrak{h}_1$ and $\mathfrak{h}_2'\mathfrak{B}\mathfrak{h}_2 = \mathfrak{h}_3'\mathfrak{B}\mathfrak{h}_3 + 2\mathfrak{h}_3'\mathfrak{B}\mathfrak{h}_1 + \mathfrak{h}_1'\mathfrak{B}\mathfrak{h}_1$. Hence, $-2\mathfrak{h}_3'\mathfrak{B}\mathfrak{h}_1 \equiv \mathfrak{h}_3'\mathfrak{B}\mathfrak{h}_3 \pmod p$ or (mod 4) according as p is odd or $p = 2$. Thus, $\mathfrak{h}_2 = \mathfrak{h}_1 + \mathfrak{h}_3 \equiv \mathfrak{h}_1 - 2\mathfrak{h}_3\mathfrak{h}_3'\mathfrak{B}\mathfrak{h}_1/(\mathfrak{h}_3'\mathfrak{B}\mathfrak{h}_3) \pmod p$ since $\mathfrak{h}_3'\mathfrak{B}\mathfrak{h}_1 = \mathfrak{h}_2'\mathfrak{B}\mathfrak{h}_1 \not\equiv 0 \pmod p$. This is true even if $p = 2$, for then $\mathfrak{h}_3'\mathfrak{B}\mathfrak{h}_1 \equiv \frac{1}{2}\mathfrak{h}_3'\mathfrak{B}\mathfrak{h}_3 \equiv 1 \pmod 2$. Then $\mathfrak{h}_2 \equiv \mathfrak{A}_0\mathfrak{h}_1 \pmod p$, where $\mathfrak{A}_0 = \mathfrak{E} - 2\mathfrak{h}_3\mathfrak{h}_3'\mathfrak{B}/(\mathfrak{h}_3'\mathfrak{B}\mathfrak{h}_3)$ is easily seen to be an automorph of $\mathfrak{B}$ and the denominators of its elements are all prime to p. Then, from Lemma 2, there is a unimodular $\mathfrak{A} \equiv \mathfrak{A}_0 \pmod q$, which therefore satisfies the conditions of this theorem.

Second, if $\mathfrak{h}_2'\mathfrak{B}\mathfrak{h}_1 \equiv 0 \pmod p$, then we can find a vector $\mathfrak{c}$ satisfying the conditions of Lemma 7 if p is odd or, with the exceptions noted, of Lemma 10 if $p = 2$. Define $\mathfrak{h}_3 = \mathfrak{h}_1 - 2\mathfrak{c}\mathfrak{c}'\mathfrak{B}\mathfrak{h}_1/(\mathfrak{c}'\mathfrak{B}\mathfrak{c})$. Then $\mathfrak{h}_3 = \mathfrak{A}_0\mathfrak{h}_1$, where $\mathfrak{A}_0 = \mathfrak{E} - 2\mathfrak{c}\mathfrak{c}'\mathfrak{B}/(\mathfrak{c}'\mathfrak{B}\mathfrak{c})$ is an automorph of $\mathfrak{B}$. Since $\mathfrak{c}'\mathfrak{B}\mathfrak{c} \not\equiv 0 \pmod p$ for p odd and $\equiv 2 \pmod 4$ for $p = 2$, there is a unimodular transformation $\mathfrak{A}_1 \equiv \mathfrak{A}_0 \pmod q$ such that $\mathfrak{h}_3 \equiv \mathfrak{A}_1\mathfrak{h}_1 \pmod q$ and $\mathfrak{A}_1'\mathfrak{B}\mathfrak{A} \equiv \mathfrak{B} \pmod q$. Then $\mathfrak{h}_3'\mathfrak{B}\mathfrak{h}_3 \equiv \mathfrak{h}_1'\mathfrak{B}\mathfrak{h}_1 \pmod q$. The first case of this proof will show the existence of a unimodular transformation $\mathfrak{A}_2$ such that $\mathfrak{h}_3 \equiv \mathfrak{A}_2\mathfrak{h}_2 \pmod p$ and $\mathfrak{A}_2'\mathfrak{B}\mathfrak{A}_2 \equiv \mathfrak{B} \pmod q$ provided $\mathfrak{h}_3'\mathfrak{B}\mathfrak{h}_2 \not\equiv 0 \pmod p$. Now

$$\mathfrak{h}_3'\mathfrak{B}\mathfrak{h}_2 \equiv \mathfrak{h}_1'\mathfrak{B}\mathfrak{h}_2 - 2\,\frac{\mathfrak{c}'\mathfrak{B}\mathfrak{h}_1}{\mathfrak{c}'\mathfrak{B}\mathfrak{c}}\,\mathfrak{c}'\mathfrak{B}\mathfrak{h}_2 \qquad \pmod p,$$

$\mathfrak{h}_1'\mathfrak{B}\mathfrak{h}_2 \equiv 0 \pmod p$, and the fact that $\mathfrak{c}$ is chosen so that $\mathfrak{c}'\mathfrak{B}\mathfrak{h}_1$ and $\mathfrak{c}'\mathfrak{B}\mathfrak{h}_2$ are prime to p shows that $\mathfrak{h}_3'\mathfrak{B}\mathfrak{h}_2 \not\equiv 0 \pmod p$. Then $\mathfrak{A}_2^{-1}\mathfrak{A}_1$ is the desired $\mathfrak{A}$.

It remains to consider the exceptions of Lemma 10.

Exception 1.2 (that is, the second pair listed under exception 1). Here we may take $\beta \equiv b_1x_1^2 + b_2x_2^2 - b_3x_3^2 - b_4x_4^2 \pmod{q}$, where $b_i \equiv 1 \pmod 4$, $i = 1, 2, 3, 4$. Now the forms

$$x_1^2 + x_2^2 - x_3^2 - x_4^2, \qquad x_1^2 + 5x_2^2 - x_3^2 - 5x_4^2,$$

$$x_1^2 + x_2^2 - x_3^2 - 5x_4^2, \qquad x_1^2 + x_2^2 - 5x_3^2 - 5x_4^2$$

have the respective automorphs $\mathfrak{A}_i$:

$$\begin{Vmatrix} 1 & 1 & 0 & 1 \\ 1 & -1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & 1 \end{Vmatrix}, \qquad \begin{Vmatrix} 1 & 5 & 0 & 5 \\ -1 & 1 & 1 & 0 \\ 0 & 5 & 1 & 5 \\ 1 & 0 & -1 & 1 \end{Vmatrix},$$

$$\begin{Vmatrix} 1 & 3 & 2 & 5 \\ 1 & -1 & -1 & 0 \\ 1 & 2 & 1 & 5 \\ 0 & 1 & 1 & 1 \end{Vmatrix}, \qquad \begin{Vmatrix} 5 & 1 & 10 & 5 \\ 1 & -5 & 5 & -10 \\ 2 & -1 & 5 & -1 \\ 1 & 2 & 1 & 5 \end{Vmatrix}.$$

Every form β above may be derived from one of these three forms by one or more of the following transformations: multiplying the form by an odd integer, permuting or interchanging the pairs (x_1, x_2) and (x_3, x_4), and stretching $x_i \to mx_i$ (m odd). The first of these transformations does not alter the automorphs; the second either leaves $\mathfrak{h}_1' = (0, 0, 1, 1)$ and $\mathfrak{h}_2' = (1, 1, 0, 0)$ unaltered or interchanges them; and the third leaves $\mathfrak{h}_1'$ and $\mathfrak{h}_2'$ (mod 2) unaltered. Then if $\mathfrak{T}$ is one of these transformations, there is a matrix $\mathfrak{A}$ with integral coefficients $\equiv \mathfrak{T}^{-1}\mathfrak{A}_i\mathfrak{T} \pmod{q}$ for the proper i, for which $\mathfrak{A}'\mathfrak{B}\mathfrak{A} \equiv \mathfrak{B} \pmod{q}$ which implies $\mathfrak{A} \equiv \pm 1 \pmod{\frac{1}{2}q}$. Furthermore, $\mathfrak{A}_i\mathfrak{h}_1 \equiv \mathfrak{h}_2 \pmod 2$ implies $\mathfrak{A}\mathfrak{h}_1 \equiv \mathfrak{h}_2 \pmod 2$ and Lemma 3 completes the proof of this case.

Exceptions 1.1 *and* 1.4. In 1.1 use the same automorphs used in the case above with a row and column (0, 0, 0, 0, 1) adjoined. Deal similarly with 1.4. In the latter, if both the vectors $\mathfrak{h}$ are replaced by their complements, the same transformation suffices. To deal with the case in which just one $\mathfrak{h}$ is replaced by its complement, write $\mathfrak{h}_3 = (0, 0, 0; 0, 1, 1)'$ and $\mathfrak{h}_1 = (1, 1, 1; 1, 0, 0)'$. We can use $\mathfrak{c} = (1, 1, 1; 0, 1, 0)'$ to satisfy the conditions of Lemma 8 and see by the first part of the proof of this theorem that there is an automorph $\mathfrak{A}_1 \pmod{q}$ such that $\mathfrak{A}_1\mathfrak{h}_3 \equiv \mathfrak{h}_1 \pmod 2$ and that there is an automorph $\mathfrak{A}_2$ such that $\mathfrak{A}_2\mathfrak{h}_3 \equiv \mathfrak{h}_2 = (1, 1, 0; 0, 0, 0)' \pmod 2$; hence the automorph $\mathfrak{A}_2\mathfrak{A}_1^{-1}$ has the property that $\mathfrak{A}_2\mathfrak{A}_1^{-1}\mathfrak{h}_1 \equiv \mathfrak{h}_2 \pmod 2$.

Exception 1.3. We here use $\mathfrak{h}_3 = (0, 0, 0; 1, 1)'$ with $\mathfrak{h}_1 = (1, 1, 1; 0, 1)'$ as above to prove the existence of an automorph $\mathfrak{A}_1$ such that $\mathfrak{A}_1\mathfrak{h}_3 \equiv \mathfrak{h}_1 \pmod 2$. Obvious modifications of the transformations used for exception 1.1 yield an

automorph $\mathfrak{A}_2$ such that $\mathfrak{A}_2\mathfrak{h}_3 \equiv \mathfrak{h}_2 = (1, 1, 0; 0, 0)'$ (mod 2) and complete the proof as above.

Exceptions 2.1 *and* 2.2. β may be taken $\equiv 2x_1x_2 + 2x_3x_4$ (mod 8) since one or both of x_2 and x_3 may be replaced by their negatives. The automorph desired is

$$\begin{Vmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{Vmatrix}.$$

Exception 2.3. Use $\mathfrak{h}_3 = (1, 1, 0, 0)'$ and proceed in a manner analogous to that for 1.3 above.

The exception of our theorem arises from the case of Lemma 10 in which $\beta \sim \beta_1 \equiv x_1^2 + \cdots + x_n^2$ (mod 2) and $\mathfrak{h}$ is a solution of $\mathfrak{x}'\mathfrak{B}_1\mathfrak{x} \equiv 0$ (mod 2) which has all its components 1. Such a solution exists if and only if n is even and, by Lemma 6b, β is equivalent to such a form as β_1 if and only if it is properly primitive. If, then, $\mathfrak{T}$ is a unimodular transformation taking β into β_1, $\mathfrak{T}\mathfrak{h}$ is a solution of $\mathfrak{x}'\mathfrak{B}\mathfrak{x} \equiv 0$ (mod 2). Since $\mathfrak{h}' = (1, \cdots, 1)$, $\mathfrak{T}\mathfrak{h}$ is the vector whose i-th component is the sum of the elements in the i-th row of $\mathfrak{T}$, for $i = 1, \cdots, n$. These components are congruent (mod 2) to the elements of the principal diagonal of $\mathfrak{T}\mathfrak{T}'$. But $\mathfrak{T}'\mathfrak{B}\mathfrak{T} \equiv \mathfrak{E}$ (mod 2) implies $\mathfrak{B}^{-1} \equiv \mathfrak{T}\mathfrak{T}'$ (mod 2) and our proof is complete.

THEOREM 3. *If p is a prime not dividing B and if $\mathfrak{h}_1$ and $\mathfrak{h}_2$ are two primitive solutions of $\mathfrak{x}'\mathfrak{B}\mathfrak{x} \equiv 0$ (mod p) while if $p = 2$ we further require that $\mathfrak{h}_1'\mathfrak{B}\mathfrak{h}_1 \equiv \mathfrak{h}_2'\mathfrak{B}\mathfrak{h}_2$ (mod 4) and that neither $\mathfrak{h}$ is of the exceptional form mentioned in Theorem 2, then $\gamma_1 \vee \gamma_2$, where γ_1 and γ_2 are the forms which β yields by use of $\mathfrak{h}_1$ and $\mathfrak{h}_2$ in the transformations $\mathfrak{J}_1$ and $\mathfrak{J}_2$, respectively.*

Proof. Since, by Theorem 1, $\mathfrak{h}$ can be replaced by any vector congruent to it (mod p) without altering the class of γ, we see that Theorem 2 shows that we may consider $\mathfrak{h}_2$ so chosen that $\mathfrak{A}\mathfrak{h}_1 \equiv \mathfrak{h}_2$ (mod q), where q is an arbitrary power of p and, for an arbitrary ∇ prime to p, $\mathfrak{h}_1 \equiv \mathfrak{h}_2$ (mod ∇). Furthermore, by Lemma 3, there is a unimodular matrix $\mathfrak{A}_3 \equiv \mathfrak{E}$ (mod ∇) and $\equiv \mathfrak{A}$ (mod q). Then $\mathfrak{A}_3\mathfrak{h}_1 \equiv \mathfrak{h}_2$ (mod $q\nabla$) and $\mathfrak{A}_3'\mathfrak{B}\mathfrak{A}_3 \equiv \mathfrak{B}$ (mod $q\nabla$). By Theorem 1, without altering the class of γ_2 we may take $\mathfrak{J}_2 = \mathfrak{A}_3\mathfrak{J}_1$ since its last column is $\mathfrak{h}_2$ and all the other columns are divisible by p. Then $p\mathfrak{C}_2 = \mathfrak{J}_2'\mathfrak{B}\mathfrak{J}_2 = \mathfrak{J}_1'\mathfrak{A}_3'\mathfrak{B}\mathfrak{A}_3\mathfrak{J}_1 \equiv \mathfrak{J}_1'\mathfrak{B}\mathfrak{J}_1 = p\mathfrak{C}_1$ (mod $q\nabla$). Hence $\mathfrak{C}_2 \equiv \mathfrak{C}_1$ (mod $q\nabla/p$) and $\gamma_2 \vee \gamma_1$ by Lemma 5, for the determinant of each form is $p^{r-2}B$ with r the number of variables and, since each is derived from $\mathfrak{B}$ by a real transformation, their indices are the same. This completes the proof of the theorem.

Before proceeding, we wish to show that the restriction $\mathfrak{h}_1'\mathfrak{B}\mathfrak{h}_1 \equiv \mathfrak{h}_2'\mathfrak{B}\mathfrak{h}_2$ (mod 4) and the ruling out of the exception in Theorem 2 are essential.

THEOREM 4. *If β is a primitive form of odd determinant and if $\mathfrak{h}_1$ and $\mathfrak{h}_2$ are proper solutions of* $\mathfrak{x}'\mathfrak{B}\mathfrak{x} \equiv 0 \pmod 2$ *such that* $\mathfrak{h}_1'\mathfrak{B}\mathfrak{h}_1 \not\equiv \mathfrak{h}_2'\mathfrak{B}\mathfrak{h}_2 \pmod 4$, *then the forms γ_1 and γ_2 corresponding to $\mathfrak{h}_1$ and $\mathfrak{h}_2$ are not of the same genus. In fact, γ_i is properly primitive if and only if* $\mathfrak{h}_i'\mathfrak{B}\mathfrak{h}_i \equiv 2 \pmod 4$.

Proof. Take $\mathfrak{h}_1'\mathfrak{B}\mathfrak{h}_1 \equiv 2 \pmod 4$ and $\mathfrak{h}_2'\mathfrak{B}\mathfrak{h}_2 \equiv 0 \pmod 4$. Then, by the first paragraph in this section,

$$\mathfrak{C}_i = \begin{pmatrix} 2\mathfrak{Z}_i^{*\prime}\mathfrak{B}\mathfrak{Z}_i^* & \mathfrak{Z}_i^{*\prime}\mathfrak{B}\mathfrak{h}_i \\ \mathfrak{h}_i'\mathfrak{B}\mathfrak{Z}_i^* & \frac{1}{2}\mathfrak{h}_i'\mathfrak{B}\mathfrak{h}_i \end{pmatrix}.$$

Now $(\mathfrak{Z}_i^*\ \mathfrak{h}_i)$ being unimodular and B odd implies

$$\begin{pmatrix} \mathfrak{Z}_i^{*\prime} \\ \mathfrak{h}_i' \end{pmatrix}\mathfrak{B}\mathfrak{h}_i = \begin{pmatrix} \mathfrak{Z}_i^{*\prime}\mathfrak{B}\mathfrak{h}_i \\ \mathfrak{h}_i'\mathfrak{B}\mathfrak{h}_i \end{pmatrix} \not\equiv \mathfrak{N} \pmod 2$$

and since $\mathfrak{h}_i'\mathfrak{B}\mathfrak{h}_i \equiv 0 \pmod 2$ we have $\mathfrak{Z}^{*\prime}\mathfrak{B}\mathfrak{h}_i \not\equiv \mathfrak{N} \pmod 2$. Hence, the γ_i are primitive forms (mod 2) and hence are primitive forms. But γ_1 is properly primitive while γ_2 is not; that is, the former but not the latter represents an odd integer. Thus no integral transformation of odd determinant can take γ_1 into a form $\equiv \gamma_2 \pmod 2$; hence they are of different genera.

THEOREM 5. *Let β be a properly primitive form of odd determinant of the type of Lemma 8 with r even and $\mathfrak{h}_1$ a vector whose r components are 1; if $\mathfrak{h}_2$ is any other primitive solution* (mod 2) *of* $\mathfrak{x}'\mathfrak{B}\mathfrak{x} \equiv 0 \pmod 2$ *and if γ_1 and γ_2 are the forms associated with $\mathfrak{h}_1$ and $\mathfrak{h}_2$, then they are not of the same genus.*

Proof. If $r = 2$, there is only one non-zero solution (mod 2) of $\mathfrak{x}'\mathfrak{B}\mathfrak{x} \equiv 0 \pmod 2$. Henceforth, we take $r \geq 4$. For the time being, we omit the subscripts on the $\mathfrak{Z}$'s, $\mathfrak{C}$'s, etc. We have $\mathfrak{Z}'\mathfrak{B}\mathfrak{Z} = 2\mathfrak{C}$. The discussion of Theorem 4 shows that $\mathfrak{C}$ is primitive. Now $\mathfrak{Z}^{-1}\mathfrak{B}^{-1}\mathfrak{Z}'^{-1} = \frac{1}{2}\mathfrak{C}^{-1}$, and since the determinant of $\mathfrak{C}$ is $2^{r-2}B$, we have $2^{r-1}B\mathfrak{Z}^{-1}\mathfrak{B}^{-1}\mathfrak{Z}'^{-1} = \operatorname{adj}\mathfrak{C}$. Noting that $\mathfrak{Z} = \mathfrak{Z}^0\mathfrak{P}$ we have

$$2^{r-1}B\mathfrak{P}^{-1}(\mathfrak{Z}^{0-1}\mathfrak{B}^{-1}\mathfrak{Z}^{0\prime\,-1})\mathfrak{P}^{-1} = \operatorname{adj}\mathfrak{C}$$

or

$$2^{r-3}B(2\mathfrak{P}^{-1}\mathfrak{Z}^{0-1})\mathfrak{B}^{-1}(\mathfrak{Z}^{0\prime\,-1}2\mathfrak{P}^{-1}) = \operatorname{adj}\mathfrak{C}.$$

Now $2\mathfrak{P}^{-1}$ is integral and all the elements of adj $\mathfrak{C}$ are divisible by 2^{r-3}. Write $\mathfrak{C}^0 = \operatorname{adj}\mathfrak{C}/2^{r-3}$ and let γ^0 be the form having this matrix. By permuting the variables in β if necessary we can take

$$\mathfrak{Z}^0 = \begin{pmatrix} \mathfrak{E} & \mathfrak{h} \\ \mathfrak{n} & 1 \end{pmatrix},$$

where $\mathfrak{h}' = (h_1, \cdots, h_{r-1})$. By virtue of the fact that $\mathfrak{B} \equiv \mathfrak{B}^{-1} \equiv \mathfrak{E} \pmod 2$, we have

$$C^0 \equiv \begin{pmatrix} \mathfrak{E} & -\mathfrak{h} \\ \mathfrak{n} & 2 \end{pmatrix} \begin{pmatrix} \mathfrak{E} & \mathfrak{n} \\ -\mathfrak{h}' & 2 \end{pmatrix}$$

$$\equiv \left\| \begin{matrix} 1 + h_1 & h_1h_2 & h_1h_3 & \cdots & h_1h_{r-1} & 0 \\ h_1h_2 & 1 + h_2 & h_2h_3 & \cdots & h_2h_{r-1} & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ h_{r-1}h_1 & h_{r-1}h_2 & \cdots & \cdots & h_{r-1} + 1 & 0 \\ 0 & 0 & \cdots & \cdots & 0 & 0 \end{matrix} \right\| \pmod 2.$$

Now γ^0 is primitive (mod 2), for, if one of the first $r - 1$ components of $\mathfrak{h}$ is even, one of the elements on the principal diagonal is odd, while if all the first $r - 1$ components of $\mathfrak{h}$ are odd, there are many odd elements off the principal diagonal. Furthermore, if $(\mathfrak{h}, 1) = \mathfrak{h}_1$, γ^0 is improperly primitive while if $(\mathfrak{h}, 1) \neq \mathfrak{h}_1$, γ^0 is properly primitive. Hence γ_1^0 and γ_2^0 are not of the same genus. This is equivalent to saying that γ_1 and γ_2 are not of the same genus.

We now proceed to consideration of *forms whose determinants* are not prime to p. For any prime p, and q, an arbitrary power of p, we may, by Lemmas 3 and 6a, consider any form

$$\text{(a)} \qquad \sigma \equiv \beta + p\lambda \pmod q,$$

where the determinant of β is prime to p and β and λ have no variables in common, β being a form in r variables and λ in $m - r$ variables. Now, in place of the transformation $\mathfrak{J}$ we use for an arbitrary ∇ (prime to p but containing sufficiently high powers of the odd factors of S and of 2 if $p \neq 2$, S being the determinant of σ) a transformation

$$\mathfrak{K} \equiv \begin{pmatrix} \mathfrak{J} & \mathfrak{N} \\ \mathfrak{N} & \mathfrak{E}^{(m-r)} \end{pmatrix} \pmod q, \qquad \mathfrak{K} \equiv \mathfrak{E} \pmod \nabla.$$

With these notations, the following theorem follows directly from Theorem 3.

THEOREM 6. *If σ is a quadratic form* (a) *above and if $\mathfrak{h}_1$ and $\mathfrak{h}_2$ satisfy the conditions of Theorem* 3 *and determine $\mathfrak{J}_1$ and $\mathfrak{J}_2$ and hence $\mathfrak{K}_1$ and $\mathfrak{K}_2$ and if $\mathfrak{K}_1$ and $\mathfrak{K}_2$ take σ into $p\omega_1$ and $p\omega_2$, respectively, then ω_1 and ω_2 are of the same genus.*

Proof. $\omega_1 \equiv \omega_2 \pmod \nabla$. Also $\omega_i \equiv \gamma_i + \lambda \pmod{q/p}$ and hence, since γ_1 and γ_2 are of the same genus, ω_1 is equivalent to a form $\equiv \omega_2 \pmod{\nabla q/p}$.

Conversely, suppose ω_1 is some form obtained from β by such a transformation and $\omega_2 \text{ v } \omega_1$. We may take $\omega_2 \equiv \omega_1 \pmod{q\nabla}$ and, applying $\mathfrak{K}^{-1}$ to $p\omega_2$ get a form $\sigma_1 \equiv \sigma \pmod{\nabla q}$ and $\sigma_1 \text{ v } \sigma$.

We now proceed to prove two theorems on the number of classes in related genera.

THEOREM 7. *If σ is a form $\equiv \beta + p\lambda$ (mod q) (see (a) above), and if $n(\sigma, p)$ is the number of distinct primitive solutions of $\mathfrak{x}'\mathfrak{B}\mathfrak{x} \equiv 0$ (mod p) such that no one is a scalar multiple of any other and no one can be obtained from any other by multiplication on the left by the matrix composed of the elements of the first r rows and columns of an automorph of σ, then the number of classes in the genus of ω (derived from σ by a transformation $\mathfrak{K}$) $\leq \sum_k n(\sigma_k, p)$, the sum being over representatives of all the classes in the genus of σ. If $p = 2$, the congruence is replaced by one of $\mathfrak{x}'\mathfrak{B}\mathfrak{x} \equiv 0$ (mod 4) or $\mathfrak{x}'\mathfrak{B}\mathfrak{x} \equiv 2$ (mod 4) with the exceptional solution $\mathfrak{h}$ of Theorem 2 barred. λ may be vacuous but β is not.*

Proof. In what follows we proceed as if λ were not vacuous. The proof is easily modified to dispose of the case when the determinant of σ is prime to p. Some transformation

$$\mathfrak{K} = \begin{pmatrix} \mathfrak{J} & \mathfrak{N} \\ \mathfrak{N} & \mathfrak{E} \end{pmatrix}$$

takes σ_1 into some $p\omega_1$. Then, for an arbitrary modulus ∇, the representatives ω_i of the other classes of the genus of ω_1 may be considered to be all congruent to ω_1 (mod ∇). $\mathfrak{K}^{-1}$ will take each of the forms $p\omega_i$ into an integral form $\sigma_i \equiv \sigma_1$ (mod ∇) and hence of the same genus as σ_1.

First, we prove that, if $\omega_2 \equiv \omega_1$ (mod ∇) and $\sigma_1 \sim \sigma_2$, where $\mathfrak{K}$ takes σ_1 into $p\omega_1$ and σ_2 into $p\omega_2$, then there is a transformation $\mathfrak{K}_1$ taking σ_1 into a form $p\omega_0 \sim p\omega_2$. Let $\mathfrak{U}$ be a unimodular transformation taking σ_1 into σ_2. Then $\mathfrak{B} = \mathfrak{K}_1^{-1}\mathfrak{U}\mathfrak{K}$ takes ω_0 into ω_2. It has determinant ± 1 and hence it remains to show that we can determine $\mathfrak{K}_1$ so that $\mathfrak{B}$ has integral elements. Write

$$\mathfrak{U} = \begin{pmatrix} \mathfrak{U}_1 & \mathfrak{U}_2 \\ \mathfrak{U}_3 & \mathfrak{U}_4 \end{pmatrix},$$

where $\mathfrak{U}_1 = \mathfrak{U}_1^{(r,r)}$, $\mathfrak{U}_4 = \mathfrak{U}_4^{(m-r,m-r)}$, etc. Since

$$\mathfrak{B} = \begin{pmatrix} \mathfrak{J}_1^{-1}\mathfrak{U}_1\mathfrak{J} & \mathfrak{J}_1^{-1}\mathfrak{U}_2 \\ \mathfrak{U}_3\mathfrak{J} & \mathfrak{U}_4 \end{pmatrix},$$

we proceed first to show that $\mathfrak{U}_2 \equiv \mathfrak{N}$ (mod p). Now

$$\mathfrak{S}_i \equiv \begin{pmatrix} \mathfrak{B}_i & \mathfrak{N} \\ \mathfrak{N} & \mathfrak{N} \end{pmatrix} \pmod{p}$$

and $\mathfrak{U}'\mathfrak{S}_1\mathfrak{U} = \mathfrak{S}_2$ implies

$$\begin{pmatrix} \mathfrak{U}_1'\mathfrak{B}_1\mathfrak{U}_1 & \mathfrak{U}_1'\mathfrak{B}_1\mathfrak{U}_2 \\ \mathfrak{U}_2'\mathfrak{B}_1\mathfrak{U}_1 & \mathfrak{U}_2'\mathfrak{B}_1\mathfrak{U}_2 \end{pmatrix} \equiv \begin{pmatrix} \mathfrak{B}_2 & \mathfrak{N} \\ \mathfrak{N} & \mathfrak{N} \end{pmatrix} \pmod{p}.$$

Also $\mathfrak{U}_1'\mathfrak{B}_1\mathfrak{U}_1 \equiv \mathfrak{B}_2 \not\equiv 0$ (mod p) implies $| \mathfrak{U}_1'\mathfrak{B}_1 | \not\equiv 0$ (mod p). Hence $\mathfrak{U}_1'\mathfrak{B}_1\mathfrak{U}_2 \equiv \mathfrak{N}$ (mod p) implies $\mathfrak{U}_2 \equiv \mathfrak{N}$ (mod p). It remains to consider $\mathfrak{J}_1^{-1}\mathfrak{U}_1\mathfrak{J}$. Here,

$\mathfrak{Z}_1^{-1} = (\mathfrak{G}_1'/p\ \mathfrak{g}_1)'$ and

$$\mathfrak{Z}_1^{-1}\mathfrak{U}_1\mathfrak{Z} = \begin{pmatrix}\mathfrak{G}_1/p \\ \mathfrak{g}_1'\end{pmatrix}\mathfrak{U}_1(\mathfrak{Z}^*p\ \mathfrak{h}) = \begin{pmatrix}\mathfrak{G}_1\mathfrak{U}_1\mathfrak{Z}^* & \mathfrak{G}_1\mathfrak{U}_1\mathfrak{h}/p \\ \mathfrak{g}_1'\mathfrak{U}_1\mathfrak{Z}^*p & \mathfrak{g}_1'\mathfrak{U}_1\mathfrak{h}\end{pmatrix}$$

and we seek to determine $\mathfrak{G}_1$ so that $\mathfrak{G}_1\mathfrak{U}_1\mathfrak{h} \equiv \mathfrak{N} \pmod{p}$. But $\mathfrak{x}'\mathfrak{U}_1\mathfrak{h} \equiv 0 \pmod{p}$ has $r - 1$ linearly independent solutions (mod p). Choose these to be the $r - 1$ rows of $\mathfrak{G}_1$ and choose $\mathfrak{g}_1$ so that $(\mathfrak{G}_1'\ \mathfrak{g}_1)'$ is unimodular. This proves $\omega_0 \sim \omega_2$.

Notice that $\omega_0 \sim \omega_1$ would imply $\omega_2 \sim \omega_1$, contrary to our choice of ω_2. Hence $\mathfrak{h}_0$ (the $\mathfrak{h}$ for ω_0) is not, by Theorem 1, a scalar multiple of the $\mathfrak{h}$ for ω_1. If we can show that for some automorph

$$\mathfrak{A} = \begin{pmatrix}\mathfrak{A}_1 & \mathfrak{A}_2 \\ \mathfrak{A}_3 & \mathfrak{A}_4\end{pmatrix}$$

of σ, where $\mathfrak{A}_1 = \mathfrak{A}_1^{(r,r)}$, etc. and $\mathfrak{h}_0 \equiv \mathfrak{A}_1\mathfrak{h} \pmod{p}$, it will follow that $\omega_0 \sim \omega_1$ and our theorem results, for in that case representatives of all the classes of the genus of ω_1 will be derived by transformations $\mathfrak{K}$ for non-equivalent forms of the genus of σ and no two $\mathfrak{h}$'s, $\mathfrak{h}_0$ and $\mathfrak{h}$, will have the property that $\mathfrak{h}_0 \equiv \mathfrak{A}_1\mathfrak{h} \pmod{p}$.

Suppose, then, $\mathfrak{h}_0 \equiv \mathfrak{A}_1\mathfrak{h} \pmod{p}$, where $\mathfrak{A}_1$ is the matrix whose elements in order are those in the first r rows and columns of an automorph of σ. Since $\mathfrak{A}_2 \equiv \mathfrak{N} \pmod{p}$, then $A_1A_4 \equiv \pm 1 \pmod{p}$. Since $\mathfrak{h}$ is primitive, we may form a matrix $(\mathfrak{T}_0\ \mathfrak{h})$ of integral elements whose determinant $\equiv A_4 \pmod{p}$. Then the matrix $(\mathfrak{A}_1\mathfrak{T}_0\ \mathfrak{A}_1\mathfrak{h})$ has integral elements and determinant $\equiv \pm 1 \pmod{p}$. Hence, by Lemma 3, there is a unimodular matrix $\mathfrak{Z}_0^0 \equiv (\mathfrak{A}_1\mathfrak{T}_0\ \mathfrak{A}_1\mathfrak{h}) \pmod{p}$. Then choose $\mathfrak{Z}_0 = \mathfrak{Z}_0^0\mathfrak{P} \equiv (p\mathfrak{A}_1\mathfrak{T}_0\ \mathfrak{A}_1\mathfrak{h}) \pmod{p}$, and the transformation

$$\mathfrak{K}_0^{-1}\mathfrak{A}\mathfrak{K} = \begin{pmatrix}\mathfrak{Z}_0^{-1}\mathfrak{A}_1\mathfrak{Z} & \mathfrak{Z}_0^{-1}\mathfrak{A}_2 \\ \mathfrak{A}_3\mathfrak{Z} & \mathfrak{A}_4\end{pmatrix}$$

takes a form in the class of ω_0 into ω_1. But $\mathfrak{A}_2 \equiv \mathfrak{N} \pmod{p}$ and the last column of $\mathfrak{A}_1\mathfrak{Z}$ is congruent (mod p) to the last column of $\mathfrak{Z}_0$. Hence, by the proof of Theorem 1, $\mathfrak{Z}_0^{-1}\mathfrak{A}_1\mathfrak{Z}$ has integral elements and we have shown that $\omega_0 \sim \omega$.

For $p = 2$, there is no change in the argument except to make the replacements indicated in the statement of the theorem.

COROLLARY. *If $r = 2$, that is, if p divides every three-rowed minor of $\mathfrak{S}$ but not every two-rowed minor, the number of classes in the genus of σ is not less than half the number of classes in the genus of ω.*

This is true because here there are at most two solutions of $\mathfrak{x}'\mathfrak{B}\mathfrak{x} \equiv 0 \pmod{p}$ yielding non-equivalent forms.

THEOREM 8. *If δ is the form obtained from σ by replacing $x_1, \cdots, x_r$ by $px_1, \cdots, px_r$ and dividing the resulting form by p, the number of classes in the genus of σ is not less than the number of classes in the genus of δ. The equality holds if the determinant of λ (where $\sigma \equiv \beta + p\lambda \pmod{q}$) is prime to p. The case when λ is vacuous is trivial and is here excluded.*

Proof. The transformation

$$\begin{Vmatrix} p\mathfrak{E} & \mathfrak{N} \\ \mathfrak{N} & \mathfrak{E} \end{Vmatrix}$$

takes σ into $p\delta$. We get just one δ from each σ of the genus and each form $\equiv \delta$ (mod ∇), for an arbitrary ∇, is derived from a form $\equiv \sigma$ (mod $p\nabla$) by the same transformation. A similar argument to that for the proof of the previous theorem shows that $\sigma_1 \sim \sigma_2$ implies $\delta_1 \sim \delta_2$.

Suppose the determinant of λ is prime to p. Then

$$\mathfrak{S} \equiv \begin{pmatrix} \mathfrak{B} & \mathfrak{N} \\ \mathfrak{N} & p\mathfrak{L} \end{pmatrix} \pmod{p^2},$$

$$p\mathfrak{D} = \begin{pmatrix} p\mathfrak{E} & \mathfrak{N} \\ \mathfrak{N} & \mathfrak{E} \end{pmatrix}\begin{pmatrix} \mathfrak{B} & \mathfrak{N} \\ \mathfrak{N} & p\mathfrak{L} \end{pmatrix}\begin{pmatrix} p\mathfrak{E} & \mathfrak{N} \\ \mathfrak{N} & \mathfrak{E} \end{pmatrix} \equiv \begin{pmatrix} \mathfrak{N} & \mathfrak{N} \\ \mathfrak{N} & p\mathfrak{L} \end{pmatrix} \pmod{p^2}.$$

If a unimodular transformation

$$\mathfrak{U} = \begin{pmatrix} \mathfrak{U}_1 & \mathfrak{U}_2 \\ \mathfrak{U}_3 & \mathfrak{U}_4 \end{pmatrix}$$

takes δ_1 into δ_2, it may be shown, as in the proof of Theorem 7, that $L \not\equiv 0$ (mod p) implies $\mathfrak{U}_3 \equiv \mathfrak{N}$ (mod p). Then

$$\begin{Vmatrix} p\mathfrak{E} & \mathfrak{N} \\ \mathfrak{N} & \mathfrak{E} \end{Vmatrix}\begin{Vmatrix} \mathfrak{U}_1 & \mathfrak{U}_2 \\ \mathfrak{U}_3 & \mathfrak{U}_4 \end{Vmatrix}\begin{Vmatrix} \mathfrak{E}/p & \mathfrak{N} \\ \mathfrak{N} & \mathfrak{E} \end{Vmatrix} = \begin{Vmatrix} \mathfrak{U}_1 & p\mathfrak{U}_2 \\ \mathfrak{U}_3/p & \mathfrak{U}_4 \end{Vmatrix}$$

is a unimodular transformation taking σ_1 into σ_2. This completes the proof.

The following example shows that the equality in the last theorem does not always hold. Let $\sigma = x_1^2 + x_2^2 + 121x_3^2$, $p = 11$, $\delta = 11x_1^2 + 11x_2^2 + 11x_3^2$. There is only one form in the genus of δ, but more than one in the genus of σ. That the equality in Theorem 7 does not necessarily hold is shown by taking $\sigma = x_1^2 + 3x_2^2 + 7x_3^2$ which is in a genus of two classes. For $p = 7$, $\mathfrak{h}$ may be taken to be $(2, 1)'$ and

$$\mathfrak{Z} = \begin{Vmatrix} 7 & 2 \\ 0 & 1 \end{Vmatrix}.$$

Hence, $\omega = 7x_1^2 + 4x_1x_2 + x_2^2 + x_3^2 \sim 3x_1^2 + x_2^2 + x_3^2$, which is in a genus of one class. The equality in Theorem 7 does not even necessarily hold if $\sigma = \beta$ as is shown for $p = 3$ by the form $\sigma = x_1^2 + 4x_2^2 + 7x_3^2$, which is in a genus of two classes. $\sum n(\sigma_k, 3)$ is 3 while the number of classes in the related genus for $p = 3$ is two. However, for $p = 7$, Theorem 8 shows that $\delta = 7x_1^2 + 28x_2^2 + x_3^2$ is in a genus of two classes.

4. **Recurrence formulas.** The general result of Siegel mentioned in the introduction specializes, when $\mathfrak{T}$ is a scalar t and σ is a form in m variables, with $m \geq 2$, into the following:

$$A_0(\mathfrak{S}, t) = \frac{\epsilon S^{-\frac{1}{2}} t^{\frac{1}{2}m-1} \pi^{\frac{1}{2}m}}{\Gamma(\frac{1}{2}m)} \prod \frac{A_q(\mathfrak{S}, t)}{q^{m-1}}, \tag{1}$$

where $\epsilon = 1$ or $\frac{1}{2}$ according as $m > 2$ or $m = 2$ and the product ranges over all $q = p^a$, where p is a prime, $a > 2b$, and b is the highest power of p dividing $2t$. If σ is a positive form $A_0(\mathfrak{S}, t) = M(\mathfrak{S}, t)/M(\mathfrak{S})$ and

$$M(\mathfrak{S}, t) = \sum_{(\mathfrak{S}_k) \in (\mathfrak{S})} \frac{A(\mathfrak{S}_k, t)}{E(\mathfrak{S}_k)}, \qquad M(\mathfrak{S}) = \sum_{(\mathfrak{S}_k) \in (\mathfrak{S})} \frac{1}{E(\mathfrak{S}_k)},$$

where the sums are over all classes of the genus (one form in each class), $A(\mathfrak{S}_k, t)$ is the number of representations of t by $\mathfrak{S}_k$ (that is, the number of solutions of $\mathfrak{x}'\mathfrak{S}_k\mathfrak{x} = t$), and $E(\mathfrak{S}_k)$ is the number of integral automorphs of $\mathfrak{S}_k$ (determinant $= \pm 1$). $A_q(\mathfrak{S}, t)$ is the number of incongruent (mod q) solutions of $\mathfrak{x}'\mathfrak{S}\mathfrak{x} \equiv t \pmod{q}$. If σ is indefinite, A_0 has a corresponding meaning. Notice that if the form is positive and there is only one class in the genus, A_0 is the number of representations of t by $\mathfrak{S}$.

For any prime p, the two genera derived from σ are represented by the forms ω and δ (see Theorem 8). We now develop a relationship among $A_0(\mathfrak{S}, pt)$, $A_0(\mathfrak{W}, t)$, and $A_0(\mathfrak{D}, t)$. Notice that, when $p = 2$, the ω may belong to any one of two or three genera, but, if p is odd, the genus of ω is determined.

Substitution in (1) shows us that

$$\begin{aligned} A_0(\mathfrak{S}, pt) &= kp^{\frac{1}{2}m-1} A_{q_1}(\mathfrak{S}, pt)/q_1^{m-1}, \\ A_0(\mathfrak{W}, t) &= kp^{1-r+\frac{1}{2}m} A_{q_2}(\mathfrak{W}, t)/q_2^{m-1}, \\ A_0(\mathfrak{D}, t) &= kp^{\frac{1}{2}m-r} A_{q_2}(\mathfrak{D}, t)/q_2^{m-1}, \end{aligned} \tag{2}$$

where k is the same for all three since $W = p^{2r-2-m}S$ and $D = p^{2r-m}S_k$ and, from the choice of $\mathfrak{R}$, the only A_q affected by our transformations is that involving our particular prime p. $q_1 = p^{a_1}$, $q_2 = p^{a_2}$, $a_1 > 2b$, $a_2 > 2(b-1)$, and b is the highest power of p dividing pt. We know [17; 542] that $A_{q_i}(\mathfrak{S}, pt)/q_i^{m-1}$ are independent of a_1 and a_2 if the inequalities involving a_1 and a_2 are satisfied. We choose $q_1 = pq$ sufficiently large and $q_2 = q$. We now need a relationship among the A_q's.

Notice first that $A_q(\mathfrak{S}, pt)$ depends only on the genus of σ since all forms in the genus may be considered to be congruent (mod ∇) for ∇ arbitrary. Below, for any vector $\mathfrak{v}$ with m components, let $\mathfrak{v}^0$ be the vector composed of its first r components. *Any vector denoted by $\mathfrak{h}$ is assumed to be primitive.*

Let $\mathfrak{h}_1$ and $\mathfrak{h}_2$ be two (primitive) solutions of $\mathfrak{x}'\mathfrak{B}\mathfrak{x} \equiv 0 \pmod{p}$, $\mathfrak{Z}_1$ and $\mathfrak{Z}_2$ the transformations, and γ_1 and γ_2 the forms associated with them. Furthermore,

suppose the $\mathfrak{h}$'s are such that $\gamma_1 \vee \gamma_2$. (They are necessarily such if p is odd.) The proof of Theorem 3 shows that, for ∇ arbitrary, $\mathfrak{J}_1$ and $\mathfrak{J}_2$ may be so chosen that $\gamma_1 \equiv \gamma_2 \pmod{q\nabla}$. Hence, we talk about just one $\mathfrak{W}$ (mod q) derived from $\mathfrak{S}$ by the transformations associated with the $\mathfrak{h}$'s if p is odd and, in addition, two other forms (representatives of genera) if $p = 2$.

Let $\mathfrak{h}$ be some fixed solution of $\mathfrak{x}'\mathfrak{B}\mathfrak{x} \equiv 0 \pmod p$ and $\mathfrak{x}_1$ a solution of

$$\mathfrak{x}'\mathfrak{S}\mathfrak{x} \equiv pt \pmod{q_1} \tag{3}$$

for which $\mathfrak{x}_1^0 \equiv s\mathfrak{h} \pmod p$, where s is a scalar prime to p. The transformation $\mathfrak{K}$ associated with $\mathfrak{h}$ yields a solution of

$$\mathfrak{y}'\mathfrak{W}\mathfrak{y} \equiv t \pmod{q} \tag{4}$$

in which, by the beginning of §3, y_r (the r-th component of $\mathfrak{y}$) $\equiv s \pmod p$. Also $\mathfrak{x}^0 = \mathfrak{J}\mathfrak{y}^0$ and the last $m - r$ components of $\mathfrak{x}$ and $\mathfrak{y}$ are identical. Furthermore, if $\mathfrak{x}_2 = \mathfrak{x}_1 + q_1\mathfrak{z}$, then $\mathfrak{x}_2^0 = \mathfrak{x}_1^0 + q_1\mathfrak{z}^0$ and $\mathfrak{y}_2^0 = \mathfrak{J}^{-1}\mathfrak{x}_2^0 = \mathfrak{y}_1^0 + \mathfrak{J}^{-1}q_1\mathfrak{z}^0 \equiv \mathfrak{y}_1^0$ (mod q). Hence, to each $\mathfrak{x}$ which is a solution of (3) corresponds exactly one solution of (4).

Conversely, for each solution of (4) with $y_r \equiv s \pmod p$ and for each primitive solution $\mathfrak{h}$ of $\mathfrak{x}'\mathfrak{B}\mathfrak{x} \equiv 0 \pmod p$ we can find a $\mathfrak{K}$ taking $\mathfrak{S}$ into $p\mathfrak{W}$, and $\mathfrak{x} = \mathfrak{K}\mathfrak{y}$ determines a solution of (3). For each solution of (3) with $y_r \equiv s \not\equiv 0 \pmod p$ the various solutions $\mathfrak{h}$ will yield distinct $\mathfrak{x}$'s since $\mathfrak{x} \equiv s\mathfrak{h} \pmod p$. There are

$$k_1 = \frac{A_p(\mathfrak{B}, 0) - 1}{p - 1}$$

distinct primitive solutions $\mathfrak{h}$, none of which is a scalar multiple of any other (mod p), where $A_p(\mathfrak{B}, 0)$ is the number of distinct solutions (mod p) of $\mathfrak{x}'\mathfrak{B}\mathfrak{x} \equiv 0$ (mod p). Thus, to each solution of (4) with $y_r \not\equiv 0 \pmod p$ correspond k_1 solutions of (3) with $\mathfrak{x}^0 \not\equiv \mathfrak{n} \pmod p$ and no $\mathfrak{x}^0$ is a scalar multiple (mod p) of any other $\mathfrak{x}^0$. We have seen that two incongruent (mod q) solutions of (4) yield incongruent solutions (mod q_1) of (3). If, on the other hand, $\mathfrak{y}_1$ is a solution of (4) with $y_r \not\equiv 0 \pmod p$ and $\mathfrak{y}_2 = \mathfrak{y}_1 + q\mathfrak{z}$, consider any primitive solution $\mathfrak{h}$ of $\mathfrak{x}'\mathfrak{B}\mathfrak{x} \equiv 0 \pmod p$. There is a corresponding transformation $\mathfrak{K}$ taking $\mathfrak{S}$ into $\mathfrak{W}$. Then $\mathfrak{x}_2^0 = \mathfrak{x}_1^0 + q\mathfrak{J}\mathfrak{z}^0 \equiv \mathfrak{x}_1^0 + q(\mathfrak{N}\,\mathfrak{h})z_r \pmod{q_1}$, where z_r is the r-th component of $\mathfrak{z}$. There are p different values for z_r (mod p) and hence p different $\mathfrak{x}^0$'s (mod q_1) for each $\mathfrak{y}_1$ (mod q). Since the last $m - r$ components of $\mathfrak{x}$ and $\mathfrak{y}$ are identical, we see that for each $\mathfrak{h}$ and each $\mathfrak{y}$ (mod q) with $y_r \not\equiv 0 \pmod p$ there are p^{m-r+1} incongruent $\mathfrak{x}$'s (mod q_1). Hence we have

$$A'_{q_1}(\mathfrak{S}, pt) = p^{m-r+1}k_1A'_q(\mathfrak{W}, t) \tag{5}$$

for p odd, where k_1 is defined above, $A'_{q_1}(\mathfrak{S}, pt)$ is the number of solutions of (3) with $\mathfrak{x}^0$ primitive, and $A'_q(\mathfrak{W}, t)$ is the number of solutions of (4) with $y_r \not\equiv 0$ (mod p). In case $p = 2$,

$$(5') \qquad A'_{e_1}(\mathfrak{S}, 2t) = 2^{m-r+1}[A'_4(\mathfrak{B}, 2)A'_e(\mathfrak{W}_1, t) + \{A'_4(\mathfrak{B}, 0) - 1\}A'_e(\mathfrak{W}_2, t) + A'_e(\mathfrak{W}_3, t)],$$

where $A'_4(\mathfrak{B}, j)$ is the number of distinct solutions (mod 2) of $\mathfrak{x}'\mathfrak{B}\mathfrak{x} \equiv j \pmod 4$, $j = 0$ or 2, excluding the exceptional solution of Theorem 2, ω_1 and ω_2 are the corresponding forms, and ω_3 the form obtained using the exceptional $\mathfrak{h}$ of Theorem 2. A'_e denotes for the respective forms the number of representations with y_r odd.

We next consider the solutions of (3) with $\mathfrak{x}^0 \equiv \mathfrak{n} \pmod p$. If $\mathfrak{x} = (x_1, x_2, \cdots, x_m)$ is such a solution of (3), $\mathfrak{z} = (x_1/p, x_2/p, \cdots, x_r/p, x_{r+1}, \cdots, x_m)$ is a solution of

$$(6) \qquad \mathfrak{x}'\mathfrak{D}\mathfrak{x} \equiv t \pmod q,$$

and conversely. For each solution of the latter there are p^{m-r} solutions of the former. Hence,

$$(7) \qquad A_{e_1}(\mathfrak{S}, pt) - A'_{e_1}(\mathfrak{S}, pt) = p^{m-r}A_e(\mathfrak{D}, t).$$

We now seek $A_e(\mathfrak{W}, t) - A'_e(\mathfrak{W}, t)$, for any form ω. If ω represents t with $y_r \equiv 0 \pmod p$, let $y_r = py_r^0$. Then, if $\mathfrak{y}_0$ is the vector composed of the first $r - 1$ components of $\mathfrak{y}$, we have

$$(p\cdot\mathfrak{J}^*\ \mathfrak{h})(\mathfrak{y}_0'\ p\cdot\mathfrak{y}_r^0)' = p(\mathfrak{J}^*\ \mathfrak{h})(\mathfrak{y}_0'\ \mathfrak{y}_r^0)'.$$

Now $p\mathfrak{C} = \mathfrak{J}'\mathfrak{B}\mathfrak{J}$ with $\mathfrak{J} = (p\cdot\mathfrak{J}^*\ \mathfrak{h})$ and if $y_r \equiv 0 \pmod p$ in $\mathfrak{y}^0$ we see that $p\mathfrak{y}^{0\prime}\mathfrak{C}\mathfrak{y}^0 = p^2\mathfrak{y}_0^{0\prime}(\mathfrak{J}^*\ \mathfrak{h})'\mathfrak{B}(\mathfrak{J}^*\ \mathfrak{h})\mathfrak{y}_0^0$, where $\mathfrak{y}_0^0$ is the same vector as $\mathfrak{y}^0$ except that its r-th component is y_r^0 instead of y_r. Thus, to every representation $\mathfrak{y}$ of t by $\mathfrak{W}$ with $y_r \equiv 0 \pmod p$ corresponds a representation $\mathfrak{z}$ of t by $\mathfrak{D}$, where the first r components of $\mathfrak{z}$ are $\mathfrak{J}^0\mathfrak{y}_0^0$ and the remaining components are the same as $\mathfrak{y}$ and $\mathfrak{J}^0 = (\mathfrak{J}^*\ \mathfrak{h})$ is unimodular. There are exactly p values of $\mathfrak{z} \pmod q$ for each $\mathfrak{y}$. Thus

$$(8) \qquad A_e(\mathfrak{D}, t) = p\{A_e(\mathfrak{W}, t) - A'_e(\mathfrak{W}, t)\}.$$

Then, using (5), (7), and (8), we get

$$(9) \qquad A_{e_1}(\mathfrak{S}, pt) = p^{m+1-r}k_1A_e(\mathfrak{W}, t) + p^{m-r}(1 - k_1)A_e(\mathfrak{D}, t) \qquad \text{for } p \text{ odd}.$$

If $p = 2$, we use (5′) instead of (5) and get

$$(10) \qquad A_{e_1}(\mathfrak{S}, 2t) = 2^{m+1-r}[A'_4(\mathfrak{B}, 2)A_e(\mathfrak{W}_1, t) + \{A'_4(\mathfrak{B}, 0) - 1\}A_e(\mathfrak{W}_2, t) + A_e(\mathfrak{W}_3, t)] + 2^{m-r}k_2A_e(\mathfrak{D}, t),$$

where $k_2 = 2 - A_2(\mathfrak{B}, 0)$.

Using (9) with (2), we get

$$(11) \qquad A_0(\mathfrak{S}, pt) = k_1A_0(\mathfrak{W}, t) + (1 - k_1)A_0(\mathfrak{D}, t) \qquad \text{for } p \text{ odd},$$

where $k_1 = \{A_p(\mathfrak{B}, 0) - 1\}/(p - 1)$.

Using (10) with (2), we get

$$
(12)\quad
\begin{aligned}
A_0(\mathfrak{S}, 2t) = A_4'(\mathfrak{B}, 2)A_0(\mathfrak{W}_1, t) &+ \{A_4'(\mathfrak{B}, 0) - 1\}A_0(\mathfrak{W}_2, t) \\
&+ A_0(\mathfrak{W}_3, t) + \{2 - A_2(\mathfrak{B}, 0)\}A_0(\mathfrak{D}, t),
\end{aligned}
$$

where the terms on the right side are defined after equation (5′).

(11) *and* (12) *are the recursion formulas which constitute the principal result of this paper.*

In case $r = m$, it follows that $\mathfrak{D} = 2\mathfrak{S}$ and (12) may be simplified. Notice that $t \equiv 1 \pmod 2$ implies $A(\mathfrak{W}_2, t) = A(\mathfrak{D}, t) = 0$ and hence $A_0(\mathfrak{W}_2, t) = A_0(\mathfrak{D}, t) = 0$. Thus, (12) becomes

$$
(12')\quad A_0(\mathfrak{S}, 2t) = A_4'(\mathfrak{B}, 2)A_0(\mathfrak{W}_1, t) + A_0(\mathfrak{W}_3, t) \qquad \text{for } r = m, t \equiv 1 \pmod 2.
$$

On the other hand, if t is even, $A_q'(\mathfrak{W}_1, t) = 0$, and hence, by (8), $A_q(\mathfrak{D}, t) = 2A_q(\mathfrak{W}_1, t)$. The genus ω_3 exists if and only if $\beta = \sigma$ is properly primitive and r is even. If the trace of $\mathfrak{B}$ is not divisible by 4, $A_q'(\mathfrak{W}_3, t) = 0$ and, by (8), $A_q(\mathfrak{D}, t) = 2A_q(\mathfrak{W}_3, t)$. Thus, if we let $w = 0$ if the genus ω_3 exists and the trace of $\mathfrak{B}$ is divisible by 4 and otherwise $w = 1$, we have

$$
A_q(\mathfrak{W}_3, t) = (1 - w)A_q(\mathfrak{W}_3, t) + A_q(\mathfrak{D}, t)w/2.
$$

Hence (10) becomes

$$
A_{q_1}(\mathfrak{S}, 2t) = 2[\{A_4'(\mathfrak{B}, 0) - 1\}A_q(\mathfrak{W}_2, t) + (1 - w)A_q(\mathfrak{W}_3, t)] + k_2'A_q(\mathfrak{D}, t),
$$

where $k_2' = 2 - A_2(\mathfrak{B}, 0) + A_4'(\mathfrak{B}, 2) + w = 2 - A_4^*(\mathfrak{B}, 0) - A_4^*(\mathfrak{B}, 2) + A_4^*(\mathfrak{B}, 2) = 2 - A_4^*(\mathfrak{B}, 0)$, where $A_4^*(\mathfrak{B}, 2a)$ is the number of distinct solutions (mod 2) of $\mathfrak{x}'\mathfrak{B}\mathfrak{x} \equiv 2a \pmod 4$. Thus, using (2), we get

$$
(12'')\quad
\begin{aligned}
A_0(\mathfrak{S}, 2t) = \{A_4'(\mathfrak{B}, 0) - 1\}A_0(\mathfrak{W}_2, t) &+ (1 - w)A_0(\mathfrak{W}_3, t) \\
&+ \{2 - A_4^*(\mathfrak{B}, 0)\}A_0(\mathfrak{S}, \tfrac{1}{2}t)
\end{aligned}
$$

when $r = m$, t is even; the genus ω_3 exists if and only if β is properly primitive and r is even; $w = 0$ if the genus of ω_3 exists and the trace of $\mathfrak{B}$ is divisible by 4; otherwise $w = 1$.

The significance of this result is more clearly seen if we notice that the determinants of ω and δ differ from that of σ only by a factor which is a power of p. Hence, by a series of such steps we can finally arrive at a t which has no factor in common with twice the determinant of $\mathfrak{S}$. Then, the evaluation of (1) is accomplished by means of the $A_q(\mathfrak{S}, t)$ for p odd and $A_8(\mathfrak{S}, t)$, where $q = p$ for any odd prime factor p of S. When p does not divide S, there are explicit formulas for $A_q(\mathfrak{S}, t)$ (see §5). If p is a factor of S, σ is equivalent to a form $\beta + p\lambda \pmod q$, where β has a determinant prime to p and has r variables; then $A_p(\mathfrak{S}, t) = p^{m-r}A_p(\mathfrak{B}, t)$. Here we have the especially simple form

$$
(13)\qquad A_p(\mathfrak{B}, t) = p^{r-1}(1 - \epsilon p^{-\frac{1}{2}r}) \qquad \text{or } p^{r-1}(1 + \epsilon p^{\frac{1}{2}(1-r)})
$$

according as r is even or odd, and $\epsilon = ((-1)^{\frac{1}{2}r}B \mid p)$ or $((-1)^{\frac{1}{2}(r-1)}Bt \mid p)$ in the respective cases. p is an odd prime not dividing Bt.

In case σ is in a genus of one class, one may easily deduce by elementary means, without using Siegel's results (1), a formula numerically equivalent to (11). It is

$$A(\mathfrak{S}, pt) = k_{11}A(\mathfrak{W}_1, t) + k_{12}A(\mathfrak{W}_2, t) + \cdots + k_{1s}A(\mathfrak{W}_s, t) + (1 - k_1)A_0(\mathfrak{D}, t), \tag{11'}$$

where ω_i are representatives of the classes of the genus related to σ and k_{1i} is the number of primitive solutions $\mathfrak{h}_i$ of $\mathfrak{x}'\mathfrak{W}\mathfrak{x} \equiv 0 \pmod p$ leading to the class ω_i, no $\mathfrak{h}_i$ being a scalar multiple of any other. This, taken with (11), suggests the interesting relationship

$$k_{1i}^{-1} = E(\mathfrak{W}_i)M(\mathfrak{W})k_1^{-1},$$

where the quantities on the right are defined in the introduction.

In the interests of more explicit formulas we proceed to give the formulas for k_1 for p odd and $A_8(\mathfrak{S}, t)$.

5. **The evaluation** of k_1 **and** $A_8(\mathfrak{S}, t)$. The value of k_1 and methods for evaluation of $A_8(\mathfrak{S}, t)$ were first given by Lebesgue and Jordan [10], [1], [8] respectively and both later by Minkowski [14; 45–58]. The methods are more accessible, however, in Siegel's work [17; 541]. We here merely state the resulting formulas.

If p is an odd prime not dividing S,

$$A_p(S, 0) = \begin{cases} p^{m-1} & \text{if } m \text{ is odd,} \\ p^{m-1} + \epsilon p^{\frac{1}{2}m-1}(p - 1) & \text{if } m \text{ is even,} \end{cases}$$

where $\epsilon = ((-1)^{\frac{1}{2}m}S \mid p)$. Hence

$$k_1 = \begin{cases} \dfrac{p^{m-1} - 1}{p - 1} & \text{if } m \text{ is odd,} \\[2ex] \dfrac{p^{m-1} - 1}{p - 1} + \epsilon p^{\frac{1}{2}m-1} & \text{if } m \text{ is even.} \end{cases} \tag{14}$$

Using Lemma 6 we see that we may consider our form to be

$$\sigma \equiv \sigma_1 + 2\sigma_2 + 4\sigma_4 \pmod 8,$$

where each σ_i is a quadratic form of odd determinant and may be taken to be of one of the forms:

$$\alpha = a_1x_1^2 + \cdots + a_rx_r^2,$$

$$\beta = 2jx_1^2 + 2bx_1x_2 + 2jx_2^2 + 2x_3x_4 + \cdots + 2x_{2r-1}x_{2r},$$

where j is 0 or 1 and b is odd. $A_8(\mathfrak{S}, t)$ may be quickly found from $A_8(\mathfrak{A}, t)$ and $A_8(\mathfrak{B}, t)$. We have

$$A_8(\mathfrak{A}, t) = 8^{r-1}[1 + 2^{-2r}c + 2^{-3r}d], \tag{15}$$

where

$$c = 2^{\frac{1}{2}(3r+s_0)}(-1)^{s_1}N, \qquad N = (1 - i)^{-s_0}i^{-t} + (1 + i)^{-s_0}i^{t},$$

s_0 is the least positive residue (mod 4) of the sum of the coefficients of α and s_1 is chosen so that $r_3 - r_1 + s_0 \equiv 4s_1 \pmod 8$, r_g being the number of coefficients of α which are congruent to g (mod 4); $d = 0, -4^{r+1}, 4^{r+1}$ according as $s_0 - t \not\equiv 0$ (mod 4), $\equiv 4$ (mod 8) or $\equiv 0$ (mod 8). Also

$$A_4(\mathfrak{A}, t) = 4^{r-1}[1 + 2^{\frac{1}{2}(s_0-r)}(-1)^{s_1}N]; \tag{16}$$

the letters used are defined below (15).

Also the following formulas hold for the second form:

$$A_8(\mathfrak{B}, 2t) = 8^{2r-1}\{2 + 2^{-r+1}(-1)^{j+t}+2^{-2r}N_t\}, \tag{17}$$

where N_t is 0, -4 or 4 according as t is odd, $\equiv 2$ (mod 4) or $\equiv 0$ (mod 4).

$$A_4(\mathfrak{B}, 2t) = 2^{4r-1}[1 + 2^{-r}(-1)^{t+j}]. \tag{18}$$

6. Applications. The above formulas and methods apply to binary forms but for such forms this machinery is not necessary. Results may be obtained just as expeditiously by more elementary means [5], [15].

We first exhibit Siegel's formulas [17; 544], [8] for $q^{1-m}A_q(\mathfrak{S}, t)$. p is a prime not dividing $2S$ and p^l is the highest power of p dividing $t = p^l t_1$. ϵ is equal to the Legendre symbol $((-1)^{\frac{1}{2}m}S \mid p)$ or $((-1)^{\frac{1}{2}(m-1)}St_1 \mid p)$, according as m is even or odd and $r = p^{1-\frac{1}{2}m}$ or p^{2-m} in the respective cases. Also $q = p^a$ and $a > l$. Then

$$\begin{aligned} q^{1-m}A_q(\mathfrak{S}, t) &= (1 - \epsilon p^{-\frac{1}{2}m})(1 + \epsilon r + \epsilon^2 r^2 + \cdots + \epsilon^l r^l) \qquad \text{for } m \text{ even}, \\ &= (1 - p^{1-m})(1 + r + r^2 + \cdots + r^{\frac{1}{2}(l-1)}) \qquad \text{for } m \text{ and } l \text{ odd}, \\ &= (1 - p^{1-m})[1 + r + r^2 + \cdots + r^{\frac{1}{2}l-1} + r^{\frac{1}{2}l}/(1 - \epsilon p^{\frac{1}{2}(1-m)})] \\ &\qquad \text{for } m \text{ odd and } l \text{ even}. \end{aligned}$$

We shall find it necessary, at times, to put subscripts on l, ϵ and r to indicate the prime concerned. In deriving explicit formulas for $A_0(\mathfrak{S}, t)$ the following lemmas are useful.

LEMMA 11. *If m is even and S and t have no odd prime factor in common,*

$$\prod_{q} q^{1-m}A_q(\mathfrak{S}, t) = \{\sum \epsilon_n n^{-\frac{1}{2}m}\}^{-1}\{\sum \epsilon_d d^{1-\frac{1}{2}m}\} \prod_{p \mid 2S} A_q(\mathfrak{S}, t)q^{1-m}, \tag{19}$$

where the first sum is over all odd positive integers n, the second over all odd divisors of t, and $\epsilon_n = ((-1)^{\frac{1}{2}m} S \mid n)$.

Proof. Siegel's formulas show that the product $A_q(S, t)q^{1-m}$ over all powers of primes not dividing $2S$ is equal to

$$\prod (1 - \epsilon_p p^{-\frac{1}{2}m}) \prod (1 + \epsilon_p r_p + \cdots + \epsilon_p^l r_p^l),$$

r and l being defined as above, the first product being over all odd primes p and the second over all odd prime divisors of t. The second product is equal to the second sum in (19) and

$$\prod_p (1 - \epsilon_p p^{-\frac{1}{2}m})^{-1} = \prod_p (1 + \epsilon_p p + (\epsilon_p p^{-\frac{1}{2}m})^2 + \cdots) = \sum \epsilon_n n^{-\frac{1}{2}m}.$$

LEMMA 12. *If m is odd and 2S and t have no factor in common,*

$$\begin{aligned} &\prod_{q \text{ odd}} A_q(\mathfrak{S}, t)q^{1-m} \\ &\qquad = (1 - 2^{1-m})^{-1} \zeta^{-1}(m-1) \{\sum_{n=1}^{\infty} \epsilon_n' n^{\frac{1}{2}(1-m)}\} \\ &\qquad\quad \times \{\sum_{v \mid t_2} v^{2-m} \prod_{p'' \mid v} [1 - ((-1)^{\frac{1}{2}(1-m)} St/v^2 \mid p'') p^{\frac{1}{2}(1-m)}]^{-1}\} \\ &\qquad\quad \times \prod_{p \mid S} A_p(\mathfrak{S}, t)(p^{m-1} - 1)^{-1}, \end{aligned} \tag{20}$$

where n is odd, $\epsilon_p' = ((-1)^{\frac{1}{2}(m-1)} St \mid p)$, *p is an odd prime, p′ denotes an odd prime factor of t occurring to an odd power, p″ an odd prime factor of t occurring to an even power and* $\zeta^{-1}(m-1)$ *is the reciprocal of the Riemann zeta function whose values for m odd are known* [22] *in terms of Bernoulli numbers.*

Proof. Write $\epsilon_p = ((-1)^{\frac{1}{2}(m-1)} St_p \mid p)$, where t_p is the largest factor of t prime to p, $r_p = p^{2-m}$ and $t = t_2^2 t_3$, where t_3 is square-free. Siegel's formulas yield for

$$\{\prod_{p \nmid S} A_q(\mathfrak{S}, t)q^{1-m}\} \prod_{p \mid S} (1 - p^{1-m})$$

the value

$$\prod_p (1 - p^{1-m}) \prod_p (1 - \epsilon_p' p^{\frac{1}{2}(1-m)})^{-1} \{\prod_{p''} (1 - \epsilon_p p^{\frac{1}{2}(1-m)})\}^{-1} L,$$

where

$$\begin{aligned} L = &\prod_{p'} (1 + r + \cdots + r^{\frac{1}{2}(l-1)}) \\ &\times \prod_{p''} \{(1 - \epsilon_p p^{\frac{1}{2}(1-m)})(1 + r + \cdots + r^{\frac{1}{2}l-1}) + r^{\frac{1}{2}l}\}. \end{aligned}$$

First, we have

$$\begin{aligned} \prod_p (1 - p^{1-m}) &= (1 - 2^{1-m})^{-1} \prod_p (1 - p^{1-m})(1 - 2^{1-m}) \\ &= (1 - 2^{1-m})^{-1} \{\zeta(m-1)\}^{-1}. \end{aligned}$$

Second, it is easily shown that

$$\prod_{p} (1 - \epsilon'_p p^{\frac{1}{2}(1-m)})^{-1} = \sum_{n=1}^{\infty} \epsilon'_n n^{\frac{1}{2}(1-m)}.$$

Third, we proceed to prove that

$$L = t_2^{2-m} \sum_{v \mid t_2} v^{m-2} \prod_{p'' \mid v} \{1 - \epsilon_p p^{\frac{1}{2}(1-m)}\}.$$

Select any set of distinct primes $p''_1, p''_2, \cdots, p''_j$ and in this paragraph denote a product over this set of primes by $\prod'$. Then, in L, the coefficient of $\prod'(-\epsilon_p p^{\frac{1}{2}(1-m)})$ is

$$\prod_{p'} (1 + r + \cdots + r^{\frac{1}{2}(l-1)}) \prod{}' (1 + r + \cdots + r^{\frac{1}{2}l-1}) \prod_{p''} r^{\frac{1}{2}l} \prod{}' r^{-\frac{1}{2}l}.$$

The coefficient of the same product on the right side of the equality to be proved is $t_2^{2-m} \sum v^{m-2}$, the sum being over all divisors v of t_2 which are multiples of the product $p''_1 p''_2 \cdots p''_j$. But

$$\sum v^{m-2} = \prod_{p'} (1 + r^{-1} + \cdots + r^{\frac{1}{2}(1-l)}) \prod{}' (r^{-1} + \cdots + r^{-\frac{1}{2}l})$$

$$= \prod_{p'} r^{\frac{1}{2}(1-l)} \prod{}' r^{-\frac{1}{2}l} \prod_{p'} (1 + r + \cdots + r^{\frac{1}{2}(l-1)}) \prod{}' (1 + r + \cdots + r^{\frac{1}{2}l-1})$$

while

$$t_2^{2-m} = \prod_{p''} r^{\frac{1}{2}l} \prod_{p'} r^{\frac{1}{2}(l-1)}.$$

Thus our result is shown.

Fourth, it remains to evaluate

$$L \prod_{p''} (1 - \epsilon_p p^{\frac{1}{2}(1-m)})^{-1} = \prod_{p''} (1 - \epsilon_p p^{\frac{1}{2}(1-m)})^{-1} t_2^{2-m} \sum_{v \mid t_2} v^{m-2} \prod_{p'' \mid v} (1 - \epsilon_p p^{\frac{1}{2}(1-m)})$$
$$= \sum_{v \mid t_2} (v/t_2)^{m-2} \prod (1 - \epsilon_p p^{\frac{1}{2}(1-m)})^{-1},$$

where the last product is over all p'' dividing t_2 but not v. If we interchange v and t_2/v, we have

$$\sum_{v \mid t_2} v^{2-m} \prod (1 - \epsilon_p p^{\frac{1}{2}(1-m)})^{-1},$$

the product being over all p'' dividing t_2 but not t_2/v. Now $\epsilon_{p''} = ((-1)^{\frac{1}{2}(m-1)} \times St_{p''} \mid p'')$. Hence $((-1)^{\frac{1}{2}(m-1)} St/v^2 \mid p'') = ((-1)^{\frac{1}{2}(m-1)} St_3 t_2^2/v^2 \mid p'') = \epsilon_{p''}$, or 0 according as p'' does not or does divide t_2/v. Hence

$$L \prod_{p''} (1 - \epsilon_p p^{\frac{1}{2}(1-m)})^{-1} = \sum_{v \mid t_2} v^{2-m} \prod_{p'' \mid v} \{1 - p^{\frac{1}{2}(1-m)} ((-1)^{\frac{1}{2}(m-1)} St/v^2 \mid p'')\}^{-1}$$

and the proof of the lemma is complete.

It does not seem to be known generally that there are finite explicit expressions for

$$\sum \epsilon_n n^{-s},$$

the sum being over all odd n, s is a positive integer and $\epsilon_n = (D \mid n)$. These have been found by H. J. S. Smith [20; 517, 518] using "methods employed by Dirichlet and Cauchy". In view of our later applications it seems worthwhile to give these formulas for $s = 1$ and 2. Substitution in Smith's formulas gives

$$\sum \epsilon_n n^{-s} = \pi^s v \prod [1 - (D_1 \mid p) p^{-2}], \tag{21}$$

where $D = D_1 D_2^2$, D_1 is square-free and different from 1, while the product is over all prime divisors of D_2 and

$$v = \begin{cases} (-1)^s s[1 - (2 \mid \Delta)/2^s]\Delta^{-\frac{1}{2}} \sum\limits_{j=1}^{\Delta} (j \mid \Delta) F_{s-1}(j/\Delta) & \text{if } D_1 \equiv 1 \pmod 4, \\ (-1)^s s(1/2\Delta^{\frac{1}{2}}) \sum\limits_{j=1,\,(j,4\Delta)=1}^{4\Delta} (D_1 \mid j) F_{s-1}(j/4\Delta) & \text{if } D_1 \not\equiv 1 \pmod 4, \end{cases}$$

$\Delta = (-1)^s D_1$, $F_0(x) = x$, $F_1(x) = \frac{1}{2}(x^2 - x)$. If $D_1 = 1$, the evaluation in terms of $\zeta(s)$ is easy.

7. **Positive ternary quadratic forms.** Since there seems to have been no explicit evaluation of $A_0(\mathfrak{S}, t)$ for all positive ternary forms we first apply the above results to get a very simple expression for $A_0(\mathfrak{S}, t)$ in terms of class number and proceed to apply the methods of this paper to other examples. This has been evaluated for several forms. See [21]. Our results can be derived from general results of H. J. S. Smith.

The formulas of Smith in the previous section may be applied to give a finite expression for $A_0(\mathfrak{S}, t)$ not involving class number, but it is not nearly as simple as that using the class number. We let $h(n)$ be the number of proper classes of properly primitive binary forms of determinant $+n$ and $F(n)$ the number of proper classes of binary forms of determinant $+n$ one of whose coefficients is odd (recall that only forms with integral matrices are considered). In the classical literature, the determinant of a binary quadratic form $ax^2 + 2bxy + cy^2$ is defined to be $b^2 - ac$. We take as the determinant $ac - b^2$. This seems a much more natural usage since it is the value of the determinant of the matrix of the form. It also conforms to the usage in forms of more than two variables. From Dirichlet [2] we have

$$h(D) = \alpha \frac{2}{\pi} D^{\frac{1}{2}} \sum (-D \mid n) n^{-1},$$

the sum being over all odd n and $\alpha = 2$ or 1 according as $D = 1$ or $\neq 1$. Thus if $D = D_1 D_2^2$, where D_1 is square-free,

$$F(D) = \sum h(D/v^2) = \frac{2D^{\frac{1}{2}}}{\pi} \sum \alpha_v v^{-1} \prod_p \{1 - (-D/v^2 \mid p)p^{-1}\}^{-1},$$

the sum being over all odd divisors v of D_2, $\alpha_v = 1$ or 2 according as $D/v^2 \neq 1$ or $=1$ and the product is over all odd primes p. Hence

$$F(D) = \frac{2D^{\frac{1}{2}}}{\pi} \sum_n (-D \mid n)n^{-1} \sum_v \alpha_v v^{-1} \prod_{p \mid v} \{1 - (-D/v^2 \mid p)p^{-1}\}^{-1},$$

the first sum being over all odd n and the second over all odd divisors v of D_2. We may write

$$F(D) = \frac{2D^{\frac{1}{2}}}{\pi} \sum_n (-D \mid n)n^{-1} \{ \sum_{v \mid D_2} v^{-1} [\prod_{p \mid v} \{1 - (-D/v^2 \mid p)p^{-1}\}^{-1}]$$
$$+ 2\epsilon D_2^{-1} \prod_{p \mid D_2} \{1 - (-1 \mid p)p^{-1}\}^{-1}\},$$

where $\epsilon = \frac{1}{2}$ or 0 according as $D_1 = 1$ or $\neq 1$. If $D_1 = 1$, the coefficient of ϵ is

$$\frac{4D_2}{\pi} \prod_p \{1 - (-D \mid p)p^{-1}\}^{-1} D_2^{-1} \prod_{p \mid D_2} \{1 - (-1 \mid p)p^{-1}\}^{-1}$$
$$= \frac{4}{\pi} \sum (-1 \mid n)n^{-1} = 1.$$

Hence

$$F(D) - \epsilon = \frac{2D^{\frac{1}{2}}}{\pi} \sum_n (-D \mid n)n^{-1} \sum_{v \mid D_2} v^{-1} \prod_{p \mid v} \{1 - (-D/v^2 \mid p)p^{-1}\}^{-1},$$

v and n being odd.

From Lemma 12, taking $D = St$ and noting that $\zeta(2) = \pi^2/6$, we have

$$\prod_{q \text{ odd}} A_q(\mathfrak{S}, t)q^{-2} = \frac{4}{\pi(St)^{\frac{1}{2}}} \tau[F(St) - \epsilon] \prod_{p \mid S} A_p(\mathfrak{S}, t)(p^2 - 1)^{-1},$$

where τ is chosen so that

$$\tau \sum_{v \mid S_2 t_2} v^{-1} \prod_{p \mid v} \{1 - (-St/v^2 \mid p)p^{-1}\}^{-1} = \sum_{v \mid t_2} v^{-1} \prod_{p \mid v} \{1 - (-St/v^2 \mid p)p^{-1}\}^{-1},$$

where $S = S_1 S_2^2$, $t = t_1 t_2^2$, S_1 and t_1 are square-free and v is odd. A little computation shows that

$$\tau^{-1} = \sum_{d \mid S_2} d^{-1} \prod_{p \mid d} \{1 - (-St/d^2 \mid p)p^{-1}\}^{-1}, \qquad \text{for } d \text{ odd},$$

the product being 1 when $d = 1$. Then, using (1), we have

$$A_0(\mathfrak{S}, t) = \frac{8}{S} [F(St) - \epsilon]\tau \prod_{p \mid S} A_p(\mathfrak{S}, t)(p^2 - 1)^{-1}\alpha_2(\mathfrak{S}, t), \tag{22}$$

where $\epsilon = \frac{1}{2}$ or 0 according as St is a perfect square or not, τ is defined above, $\alpha_2(\mathfrak{S}, t)$ is $A_8(\mathfrak{S}, t)/8^2$ and $A_q(\mathfrak{S}, t)$ is the number of solutions of the congruence $\mathfrak{x}'\mathfrak{S}\mathfrak{x} \equiv t \pmod q$. Notice that St is a perfect square if and only if S and t are.

Example 1. Let $S = 1$. There is only one class of forms of determinant 1, which is represented by $\sigma = x_1^2 + x_2^2 + x_3^2$ and $A_0(\mathfrak{S}, t)$ is the number of representations of t by σ. (22) and (15) yield

$$A(\mathfrak{S}, t) = 8\{F(t) - \epsilon\}\lambda, \qquad \text{for } t \text{ odd}, \tag{23}$$

where $\epsilon = \frac{1}{2}$ or 0 according as t is a perfect square or not and $\lambda = 3/2$, 1 or 0 according as $t \equiv 1 \pmod 4$, $\equiv 3 \pmod 8$ or $\equiv 7 \pmod 8$. This is equivalent to a result of Kronecker's [9]. If t $\equiv 0 \pmod 4$, it is easily seen that x_1, x_2, x_3 are all even and

$$A(\mathfrak{S}, 4t) = A(\mathfrak{S}, t), \tag{23$'$}$$

while $A(\mathfrak{S}, t) = 3A(\mathfrak{S}_1, \frac{1}{2}t)$ if $t \equiv 2 \pmod 4$, where $\sigma_1 = x_1^2 + x_2^2 + 2x_3^2$. Then

$$A(\mathfrak{S}, t) = 3A(\mathfrak{S}_1, \tfrac{1}{2}t) = 12F(t) \qquad \text{if } t \equiv 2 \pmod 4. \tag{24}$$

Example 2. Application of the methods of this paper to the above forms yields other results of interest. If p is any odd prime, let ω_p be the genus of forms derived from σ of Example 1 by a transformation J. (11) yields

$$A(\mathfrak{S}, pt) = k_1 A_0(\mathfrak{W}_p, t) + (1 - k_1)A(p\mathfrak{S}, t) \tag{25}$$

and $A(p\mathfrak{S}, t) = 0$ or $A(\mathfrak{S}, t/p)$ according as $t \not\equiv 0 \pmod p$ or $\equiv 0 \pmod p$. By (14), $k_1 = p + 1$. Notice that in (25) t is not necessarily prime to p although p is a factor of the determinant of $\mathfrak{W}_p$.

Suppose $p = 7$. A solution of $\mathfrak{x}'\mathfrak{S}\mathfrak{x} \equiv 0 \pmod 7$ is $(3, 2, 1)'$ and taking

$$J = \begin{Vmatrix} 7 & 0 & 3 \\ 0 & 7 & 2 \\ 0 & 0 & 1 \end{Vmatrix}$$

we get the form $\omega = \gamma = 7x_1^2 + 7x_2^2 + 2x_3^2 + 6x_1x_3 + 4x_2x_3$. The unimodular transformation

$$\begin{Vmatrix} 0 & -1 & 1 \\ 0 & -1 & 0 \\ 1 & 3 & -1 \end{Vmatrix}$$

takes γ into $2x_1^2 + 2x_2^2 + 3x_3^2 + 2x_1x_2 + 2x_2x_3 + 2x_1x_3$. Hence

$$A_0(\mathfrak{W}_p, t) = \{A(\mathfrak{S}, tp) + pA(\mathfrak{S}, t/p)\}/(p + 1), \qquad \text{with } p = 7, \tag{26}$$

where $A(\mathfrak{S}, t)$ is given by (23), (23$'$), (24) and $A(\mathfrak{S}, t/p)$ is taken to be zero if $t \not\equiv 0 \pmod p$. Furthermore, all the primitive solutions of $\mathfrak{x}'\mathfrak{S}\mathfrak{x} \equiv 0 \pmod 7$

are scalar multiples of the solution given or may be obtained from it by permuting the components and/or changing the signs of some of them. Since so permuting or changing the signs of the components of $\mathfrak{h}$ is equivalent to multiplying it on the left by an automorph of σ we see by Theorem 7 that ω_7 is in a genus of one class and hence $A_0(\mathfrak{W}_7, t)$ is the number of representations of t by ω_7 and, in (26), the subscript on A may be omitted for $p = 7$.

Similarly, using the same σ and $p = 3, 5, 11$ we have (26) for $\omega_3, \omega_5, \omega_{11}$, respectively, equal to $x_1^2 + 2x_2^2 + 2x_3^2 - 2x_2x_3$, $x_1^2 + x_2^2 + 5x_3^2$ and $x_1^2 + 2x_2^2 + 6x_3^2 - 2x_2x_3$, all of which are in genera of one class, by the above argument.

We have a different type of result for $p = 13$. Here there are two vector solutions $\mathfrak{h}$: (4, 3, 1) and (5, 0, 1). Neither is a scalar multiple of the other or a multiple of the other by an automorph but all other solutions may be so obtained from these two. Hence there are, by Theorem 7, at most two classes in the genus of ω_{13}. In fact, there are two which are represented by $\omega_{13} = 2x_1^2 + 2x_2^2 + 5x_3^2 + 2x_1x_3 + 2x_1x_2 + 2x_2x_3$ and $\omega_{13}' = x_1^2 + x_2^2 + 13x_3^2$. We have (31) for $p = 13$, where now

$$A_0(\mathfrak{W}_{13}, t) = \frac{1}{7}\{4A(\mathfrak{W}_{13}, t) + 3A(\mathfrak{W}_{13}', t)\}$$

since ω_{13} has 12 automorphs (double the number of automorphs usually given [7] since the determinant may be -1 as well as $+1$) and ω_{13}' has 16 automorphs.

Example 3. To see how the methods apply to a form with cross products, take $\omega_7 = 2x_1^2 + 2x_2^2 + 3x_3^2 + 2x_1x_2 + 2x_1x_3 + 2x_2x_3$ as a new σ and $p = 3$. It is easiest to find $\mathfrak{h}$ directly in this case but we shall illustrate the general method. To put σ in canonical form, add the first column of its matrix to the second and third and then add the third column to the second; then perform the same operations on the columns. This is equivalent to transforming the form by

$$\mathfrak{T} = \begin{Vmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{Vmatrix}$$

and yields $\sigma_1 = 2x_1^2 + 23x_2^2 + 7x_3^2 + 12x_1x_2 + 6x_1x_3 + 24x_2x_3 \equiv 2x_1^2 + 2x_2^2 + x_3^2 \pmod 3$, and the values of $\mathfrak{h}$ may be seen by inspection to be $(1, 0, \pm 1)$, $(0, 1, \pm 1)$ together with scalar multiples of these. Multiplying these on the left by $\mathfrak{T}$ yields

$$\mathfrak{h}_1 = (-1, 0, 1), \quad \mathfrak{h}_2 = (0, 0, -1), \quad \mathfrak{h}_3 = (0, 1, -1), \quad \mathfrak{h}_4 = (1, 1, 0)$$

as values of $\mathfrak{h}$ for the form σ. Now the automorphs

$$\mathfrak{A}_1 = \begin{Vmatrix} 0 & -1 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & -1 \end{Vmatrix}, \quad \mathfrak{A}_2 = \begin{Vmatrix} 1 & 0 & 0 \\ -1 & -1 & -1 \\ 0 & 0 & 1 \end{Vmatrix}$$

have the property that $\mathfrak{A}_1\mathfrak{b}_1 = \mathfrak{b}_2$ and $\mathfrak{A}_2\mathfrak{b}_3 = \mathfrak{b}_2$. Hence, by Theorem 7, there are no more than two classes in the genus of ω. $\mathfrak{b}_1$ yields a form equivalent to $\omega_{73} = x_1^2 + 5x_2^2 + 5x_3^2 - 4x_2x_3$ and $\mathfrak{b}_4$ yields a form equivalent to $\omega_{73}' = 2x_1^2 + 2x_2^2 + 7x_3^2 - 2x_1x_2$. Hence

$$(27) \qquad A_0(\mathfrak{W}_{73}, t) = \tfrac{1}{4}\{A(\mathfrak{W}_7, t) + 3A(\mathfrak{W}_7, t/3)\},$$

where $A(\mathfrak{W}_7, t/3)$ is taken to be zero if $t \not\equiv 0 \pmod 3$ and, since ω_{73} and ω_{73}' have 8 and 24 automorphs respectively,

$$A_0(\mathfrak{W}_{73}, t) = \{3A(\mathfrak{W}_{73}, t) + A(\mathfrak{W}_{73}', t)\}/4.$$

Example 4. We finally consider a case in which t and S have a factor in common. Choose as σ the form $x_1^2 + 2x_2^2 + 5x_3^2$ in a genus of one class. If t is prime to 10, the formulas yield [21; 59]

$$(28) \qquad A(\mathfrak{S}, t) = F(10t).$$

Now $\sigma = 5t$ implies $x_1 \equiv x_2 \equiv 0 \pmod 5$ and $A(\mathfrak{S}, 5t) = A(\mathfrak{S}_1, t)$, where $\sigma_1 = 5x_1^2 + 10x_2^2 + x_3^2$ since, by Theorem 8, σ_1 is in a genus of one class. Then

$$(28') \qquad A(\mathfrak{S}, 5t) = A(\mathfrak{S}_1, t) = F(50t)[1 + (t \mid 5)]/7, \qquad \text{for } t \text{ prime to } 10,$$

and

$$(28'') \qquad A(\mathfrak{S}, 25t) = A(\mathfrak{S}, t).$$

Consider t even. The form β of the theory is $x_1^2 + 5x_3^2$, $A_4'(\mathfrak{B}, 2) = 0$, $A_4'(\mathfrak{B}, 0) = 1$ and γ_3 is the form $2x_1^2 + 2x_1x_3 + 3x_3^2$. Hence, (12) yields

$$A(\mathfrak{S}, 2t) = A(\mathfrak{W}_3, t),$$

where $\omega_3 = 2x_1^2 + x_2^2 + 3x_3^2 + 2x_1x_3$, whose representations may be found for t odd.

8. **Positive quaternary forms.** The number of representations function for various quaternaries was given by Liouville [11], [12]. A few other forms are dealt with in neighboring volumes of his Journal. From (1) and Lemma 11 we find, for t prime to $2S$,

$$(29) \qquad A_0(\mathfrak{S}, t) = S^{-\frac{1}{2}} t\pi^2 \{\textstyle\sum \epsilon_n n^{-2}\}^{-1} \{\textstyle\sum \epsilon_d d^{-1}\} \displaystyle\prod_{p \mid 2S} A_q(\mathfrak{S}, t) q^{-3},$$

where the first sum is over all odd positive n, the second over all divisors of t and $\epsilon_n = (S \mid n)$.

If S is a perfect square, $\sum \epsilon_n n^{-2} = \sum n^{-2} = \frac{3}{4}\zeta(2) = \pi^2/8$ and $t \sum \epsilon_d d^{-1} = \sum d$, that is, the sum of the divisors of t. Hence,

$$A_0(\mathfrak{S}, t) = 8S^{-1}\{\sum_{d \mid t} d\} \prod_{p \mid 2S} A_q(\mathfrak{S}, t)q^{-3}. \tag{30}$$

If S is not a perfect square, we use (21). Since there is no essential simplification, there is no great point in exhibiting the complete formula in one piece.

Example 1. If $S = 1$, we have the usual formula [6] for the number of representations of t by the sum of four squares

$$A_0(\mathfrak{S}, t) = A(\mathfrak{S}, t) = 8 \sum_{d \mid t} d, \qquad \text{for } t \text{ odd}. \tag{31}$$

Example 2. If $\sigma = x_1^2 + x_2^2 + 2x_3^2 + 2x_4^2$, we have [13]

$$A_0(\mathfrak{S}, t) = A(\mathfrak{S}, t) = 4 \sum_{d \mid t} d, \qquad \text{for } t \text{ odd}, \tag{32}$$

since Theorem 7 may be used to show that σ is in a genus of one class because the sum of four squares is. If we take $p = 3$, we see that the only two primitive solutions $\mathfrak{y}$ of $\mathfrak{x}'\mathfrak{S}\mathfrak{x} \equiv 0 \pmod 3$ which need be considered in getting the classes of ω are $\mathfrak{y}_1 : (0, 1, 0, 1)$ and $\mathfrak{y}_2 : (1, 1, 1, 1)$. They yield the forms $\omega_1' = 3x_1^2 + 3x_2^2 + 6x_3^2 + x_4^2 + 2x_2x_4 \sim \omega_1 = x_1^2 + 2x_2^2 + 3x_3^2 + 6x_4^2$ and $\omega_2 = 3x_1^2 + 3x_2^2 + 6x_3^2 + 2x_4^2 + 2x_1x_4 + 2x_2x_4 + 4x_3x_4$. Theorem 7 shows that there are not more than two classes in the genus of ω_1. ω_2, having the minimum 2, is not in the same class as ω_1. Hence, there are exactly two classes in the genus of ω_1 and their representatives are ω_1 and ω_2. Now k_1 is, by (14), equal to 16. Hence, from (11) we have

$$A_0(\mathfrak{S}, 3t) = 16A_0(\mathfrak{W}, t) - 15A_0(3\mathfrak{S}, t)$$

and (32) implies

$$16A_0(\mathfrak{W}, t) = 4 \sum_{d \mid 3t} d + 60 \sum_{3d \mid t} d, \qquad \text{for } t \text{ odd}, \tag{33}$$

the second sum being omitted if $t \not\equiv 0 \pmod 3$ and

$$2A_0(\mathfrak{W}, t) = A(\mathfrak{W}_1, t) + A(\mathfrak{W}_2, t)$$

since ω_1 and ω_2 each have 16 automorphs. The form ω_1 has been considered on numerous occasions [4]. The fact that it occurs in a genus of two classes shows that one should expect to have difficulty in finding the number of representations of any number by it. (33) may be deduced by elementary means. The formula for the number of representations of an even number by ω_1 has been found.

Example 3. Let $\sigma = x_1^2 + x_2^2 + x_3^2 + 3x_4^2$. Formula (21) yields

$$\sum (3 \mid n)n^{-2} = \frac{\pi^2}{2 \cdot 3^{\frac{1}{2}}} \sum (3 \mid j)\{(j/12)^2 - (j/12)\},$$

the sum being over $j = 1, 5, 7, 11$. Now

$$\sum (3 \mid j)\{(j/12)^2 - (j/12)\} = 1/3.$$

Hence, from (29),

$$A(\mathfrak{S}, t) = A_0(\mathfrak{S}, t) = 6 \sum_{d \mid t} (3 \mid d)t/d \prod_{p=2,3} A_q(\mathfrak{S}, t)q^{-3},$$

for t prime to 6. Now $A_3(\mathfrak{S}, t) = 3A_3(\mathfrak{S}_1, t)$, where $\sigma_1 = x_1^2 + x_2^2 + x_3^2$. Hence $A_3(\mathfrak{S}, t) = 3^3\{1 + (-t \mid 3)/3\}$ and, by (15), $A_8(\mathfrak{S}, t) = 8^3\{1 - i^{t+1}/2\}$ and [3], [12]

$$(34) \qquad A(\mathfrak{S}, t) = 6 \sum_{d \mid t} (3 \mid d)td^{-1}\{1 + (-t \mid 3)/3\}\{1 - i^{t+1}/2\}, \qquad t \text{ prime to } 6.$$

The odd multiples of 3 represented by σ may be dealt with in a manner similar to that used for ω_3 in Example 2 of §7.

Since the development for even values of t illustrates our methods, we give it here in detail. Essentially the only solutions of $\mathfrak{x}'\mathfrak{S}\mathfrak{x} \equiv 2 \pmod 4$ are $\mathfrak{h}_1 = (1, 1, 0, 0)$ and $\mathfrak{h}_3 = (1, 1, 1, 1)$, while essentially the only primitive solution of $\mathfrak{x}'\mathfrak{S}\mathfrak{x} \equiv 0 \pmod 4$ is $\mathfrak{h}_2 = (0, 0, 1, 1)$. The corresponding forms are equivalent to

$$\omega_1 = x_1^2 + x_2^2 + 2x_3^2 + 6x_4^2,$$
$$\omega_2 = 2x_1^2 + 2x_2^2 + 2x_3^2 + 2x_4^2 - 2x_1x_2,$$
$$\omega_3 = 2x_1^2 + 2x_2^2 + 2x_3^2 + 3x_4^2 - 2x_1x_4 - 2x_2x_4 - 2x_3x_4.$$

Hence, by (12),

$$(35) \qquad A_0(\mathfrak{S}, 2t) = A_4'(\mathfrak{B}, 2)A_0(\mathfrak{W}_1, t) + \{A_4'(\mathfrak{B}, 0) - 1\}A_0(\mathfrak{W}_2, t) + A_0(\mathfrak{W}_3, t) - 6A_0(2\mathfrak{S}, t).$$

It is to be noted that, since, by Theorem 7, all the forms are in genera of one class, the subscript of A may be omitted. Furthermore, from (16), $A_4'(\mathfrak{B}, 2) = 3$ and, from the discussion after (12), we have

$$(35') \qquad A(\mathfrak{S}, 2t) = 3A(\mathfrak{W}_1, t) + A(\mathfrak{W}_3, t), \qquad \text{if } t \equiv 1 \pmod 2,$$

$$(35'') \qquad A(\mathfrak{S}, 2t) = 3A(\mathfrak{W}_2, t) - 2A(\mathfrak{S}, \tfrac{1}{2}t), \qquad \text{if } t \equiv 0 \pmod 2.$$

$A(\mathfrak{W}_1, t)$ and $A(\mathfrak{W}_3, t)$ for $t \equiv 1 \pmod 2$ may be found by similar methods to that used for $A(\mathfrak{S}, t)$. It remains to consider $A(\mathfrak{W}_2, t)$ for $t \equiv 0 \pmod 2$. In ω_2 the β of our theory is $2x_1^2 - 2x_1x_2 + 2x_2^2$. The primitive solutions of $\mathfrak{x}'\mathfrak{B}\mathfrak{x} \equiv 0 \pmod 2$ are $(0, 1)$, $(1, 0)$ and $(1, 1)$, where automorphs of β take each of the first two into the last. Hence, if ω_2 is taken as our new σ, the related ω (call it ω_{21}) is $\frac{1}{2}\{2(2x_1 + x_2)^2 - 2(2x_1 + x_2)x_2 + 2x_2^2 + 2x_3^2 + 2x_4^2\}$. That is, $\omega_{21} = 4x_1^2 + x_2^2 + 2x_1x_2 + x_3^2 + x_4^2 \sim \sigma$. The form δ for ω_2 is $\delta_1 = 4x_1^2 - 4x_1x_2 + 4x_2^2 + x_3^2 + x_4^2$. $A(\mathfrak{W}_2, t) = 3A(\mathfrak{S}, \frac{1}{2}t) - 2A(\mathfrak{D}_1, \frac{1}{2}t)$. Hence (35″) becomes

$$(35''') \qquad A(\mathfrak{S}, 2t) = 7A(\mathfrak{S}, \tfrac{1}{2}t) - 6A(\mathfrak{D}_1, \tfrac{1}{2}t), \qquad \text{for } t \text{ even}.$$

If $\frac{1}{2}t$ is even, the process continues.

Bibliography

1. P. Bachmann, *Die Arithmetik der Quadratischen Formen*, vol. I, Leipzig, 1898, p. 491.
2. P. G. Lejeune-Dirichlet, *Vorlesungen über Zahlentheorie*, edited by Dedekind, 4th revision, 1894, p. 243.
3. G. Eisenstein, *Neue Theoreme der höheren Arithmetik*, Journal für Mathematik, vol. 35(1847), p. 134.
4. L. W. Griffiths, *Representation of integers in the form* $x^2 + 2y^2 + 3z^2 + 6w^2$, American Journal of Mathematics, vol. 51(1929), pp. 61–66.
5. N. A. Hall, *The number of representations function for binary quadratic forms*, American Journal of Mathematics, vol. 62(1940), pp. 589–598.
6. C. G. J. Jacobi, *Werke* I, Berlin, 1881, pp. 423–424.
7. B. W. Jones, *Report of the committee on tables of positive ternary quadratic forms*, Bulletin of the National Research Council, no. 97, 1935.
8. C. Jordan, *Sur la forme canonique des congruences du second degré et le nombre de leurs solutions*, Journal de Mathématiques, 2nd series, vol. 17(1872), pp. 368–402.
9. L. Kronecker, *Ueber die Anzahl der verschiedenen Classen quadratischer Formen von negativer Determinante*, Journal für Mathematik, vol. 57(1860), p. 253.
10. M. Lebesgue, *Recherches sur les nombres*, Journal de Mathématiques, (2), vol. 2(1837), pp. 266–275.
11. J. Liouville, Journal de Mathématiques, (2), vol. 7(1862).
12. J. Liouville, *Sur la forme* $x^2 + y^2 + z^2 + 3t^2$, Journal de Mathématiques, (2), vol. 8(1863), pp. 105–114.
13. J. Liouville, *Sur la forme* $x^2 + y^2 + 2(z^2 + t^2)$, Journal de Mathématiques, 2nd series, vol. 5(1860), pp. 269–272.
14. H. Minkowski, *Gesammelte Abhandlungen*, I, Leipzig and Berlin, 1911.
15. Gordon Pall, *Binary quadratic discriminants differing by square factors*, American Journal of Mathematics, vol. 57(1935), pp. 789–799.
16. Gordon Pall, *On the order invariants of integral quadratic forms*, Quarterly Journal of Mathematics, vol. 6(1935), pp. 35–38.
17. C. L. Siegel, *Ueber die analytische Theorie der quadratischen Formen*, Annals of Mathematics, vol. 36(1935), pp. 527–606.
18. C. L. Siegel, *Ueber die analytische Theorie der quadratischen Formen*, II, Annals of Mathematics, vol. 37(1936), pp. 230–263.
19. C. L. Siegel, *Equivalence of quadratic forms*, American Journal of Mathematics, vol. 63 (1941), pp. 658–680.
20. H. J. S. Smith, *Collected Mathematical Papers*, vol. I, Oxford, 1894.
21. J. V. Uspensky, *On the number of representations of integers by certain ternary quadratic forms*, American Journal of Mathematics, vol. 51(1929), pp. 51–60.
22. E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis*, 3rd edition, Cambridge, 1920, p. 269.

Cornell University.

A PROPERTY OF PSEUDO-CONFORMAL TRANSFORMATIONS IN THE NEIGHBORHOOD OF BOUNDARY POINTS

By Stefan Bergman and D. C. Spencer

1. There are roughly two kinds of theorems in the theory of functions of two complex variables: (i) theorems concerning the behavior of a single function of two variables; (ii) theorems involving the simultaneous behavior of a pair of functions whose Jacobian does not vanish identically. Such pairs $w_k = f_k(z_1, z_2)$, $k = 1, 2$, define mappings $\mathbf{w}$ of four-dimensional domains of the (z_1, z_2)-space into domains of the (w_1, w_2)-space; we call these mappings PT's (pseudo-conformal transformations).

In some cases the investigation of PT's may be reduced to the study of a single function of two variables (see, for example, [1]; also [8] where, by using the theory of distinguished boundary surfaces, a generalization of Lindelöf's convergence theorem is obtained); but in others reduction to one function is impossible. The problem considered here is of the latter type, and concerns geometrical properties of the boundary sets.

2. We denote manifolds by German letters, the superscript denoting the dimension of the manifold. However, we sometimes omit the superscript if the manifold is four-dimensional or if it is a point. In operations involving sets we use the customary symbols. For example, $\mathfrak{C}_1^m \cap \mathfrak{C}_2^n$ denotes the intersection of $\mathfrak{C}_1^m$ and $\mathfrak{C}_2^n$, $\mathfrak{a}^m \,\varepsilon\, \mathfrak{C}^n$ means that $\mathfrak{a}^m$ is an element of $\mathfrak{C}^n$, and the symbol $\mathfrak{S}$ denotes logical summation over a set. We write $E[\ \ ]$ for the set of points whose coordinates satisfy the conditions indicated in the bracket.

3. The theorem in one variable which we generalize may be formulated as follows. (See [7; 434]. The form of statement below conforms to that chosen in two variables.) To avoid repetition we suppose throughout that k assumes the values 1 and 2.

Let $\mathfrak{s}^1(r)$, $0 < r < r_0$, be an open arc of the circle $E[\,|z| = r]$. We denote the end-points of $\mathfrak{s}^1(r)$ by $\mathfrak{b}_1^0(r)$, $\mathfrak{b}_2^0(r)$, and write

$$\mathfrak{N}^2 = \underset{0<r<r_0}{\mathfrak{S}} \mathfrak{s}^1(r). \tag{3.1}$$

Suppose that the function $f(z)$ is regular in $\mathfrak{N}^2$; let $\mathfrak{l}_k(r)$ be the set of limit values of f as z approaches $\mathfrak{b}_k^{(0)}(r)$ through $\mathfrak{s}^1(r)$, and let $\mathfrak{C}_k$ be the aggregate of limit points of $\mathfrak{l}_k(r)$ as $r \to 0$. If

$$\iint_{\mathfrak{N}^2} |f'|^2 \rho \, d\rho \, d\varphi < \infty \qquad (z = \rho e^{i\varphi}), \tag{3.2}$$

Received January 27, 1942.

then

$$\mathfrak{C}_1 \cap \mathfrak{C}_2 \neq 0. \tag{3.3}$$

This result is intuitive when f is schlicht and bounded. For, if $\mathfrak{C}_1 \cap \mathfrak{C}_2 = 0$, the boundary of the domain $\mathfrak{N}^2$ would have a "gap" in it of width equal to the distance between $\mathfrak{C}_1$ and $\mathfrak{C}_2$. In the general case the alternative is that the values of the function "spill" through the gap, so to say, and create infinite area. We now state an analogous result for PT's.

4. Let $\mathfrak{P}^2$ be a bounded domain of $E[r_k > 0]$ for which

$$\iint_{\mathfrak{P}^2} \frac{dr_1 dr_2}{r_1 r_2} = \infty. \tag{4.1}$$

Then $\mathfrak{P}^2$ has at least one boundary point $\mathfrak{p}$ on one of the lines $r_k = 0$, and we suppose for simplicity that $\mathfrak{p}$ is the only such boundary point. For each $(r_1, r_2) \varepsilon \mathfrak{P}^2$, let $\mathfrak{M}^2(r_1, r_2)$ be a simply-connected (open) sub-domain of the square $E[0 \leq \psi_k < 2\pi]$, and write

$$\mathfrak{S}_1^2(r_1, r_2) = E\,[z_1 = r_1 e^{i\psi_1}, z_2 = r_2\, e^{i\psi_2}, (\psi_1, \psi_2) \,\varepsilon\, \mathfrak{M}^2(r_1, r_2)]. \tag{4.2}$$

We suppose that the boundary $\mathfrak{b}_1^1(r_1, r_2)$ of $\mathfrak{S}_1^2(r_1, r_2)$ is a Jordan curve defined by $z_k = \tau_k(t) = \tau_k(r_1, r_2; t)$, $0 \leq t \leq 1$, $\tau_k(0) = \tau_k(1)$, and write

$$\mathfrak{b}_1^1(r_1, r_2) = \mathop{S}_{\gamma=1}^{4} \mathfrak{b}_{1,\gamma}^1(r_1, r_2), \tag{4.3}$$

where

$$\begin{aligned} &\mathfrak{b}_{1,\gamma}^1(r_1, r_2) = E\,[z_k = \tau_k(t),\ t_{\gamma-1} \leq t \leq t_\gamma], \\ &t_0 = 0,\ t_4 = 1,\ t_{\gamma-1} < t_\gamma,\ t_\gamma = t_\gamma(r_1, r_2). \end{aligned} \tag{4.4}$$

That is to say, we make (for each $(r_1, r_2) \varepsilon \mathfrak{P}^2$) an arbitrary dissection of $\mathfrak{b}_1^1(r_1, r_2)$ into four parts. We suppose that the functions $f_k(z_1, z_2)$, $z_k = r_k e^{i\psi_k}$ are regular in the domain

$$\mathfrak{N} = \mathop{S}_{(r_1, r_2)\, \varepsilon\, \mathfrak{P}^2} \mathfrak{S}_1^2(r_1, r_2), \tag{4.5}$$

and that they define a PT $\mathbf{w}$ which is one-one and continuous (topological) in the closure of $\mathfrak{S}_1^2(r_1, r_2)$, $(r_1, r_2) \varepsilon \mathfrak{P}^2$. Finally, writing

$$J(z_1, z_2) = \left|\frac{\partial(f_1, f_2)}{\partial(z_1, z_2)}\right|^2, \qquad H(z_1, z_2) = \left|\sum_{k=1}^{2} \frac{\partial f_k}{\partial z_1} \frac{\partial f_k^*}{\partial z_2^*}\right|^2 \tag{4.6}$$

(* denotes the conjugate), we suppose that

$$\iiiint_{\mathfrak{R}} [J + H]\, d\omega < \infty, \tag{4.7}$$

where $d\omega$ denotes an element of volume of $\mathfrak{R}$.

Let $\mathfrak{b}^1_{2.\gamma}(r_1, r_2)$ be the portion of the boundary of $\mathbf{w}[\mathfrak{S}^2_1(r_1, r_2)]$ corresponding to $\mathfrak{b}^1_{1.\gamma}(r_1, r_2)$, and let $\mathfrak{C}_\gamma$ be the set of limit points of $\mathfrak{b}^1_{2.\gamma}(r_1, r_2)$ as the point (r_1, r_2) tends to $\mathfrak{p}$ in all possible ways through $\mathfrak{P}^2$.

THEOREM. *One of the two intersections $\mathfrak{C}_1 \cap \mathfrak{C}_3$, $\mathfrak{C}_2 \cap \mathfrak{C}_4$ is non-null.*

The condition that $\mathbf{w}$ is schlicht and continuous in the closure of $\mathfrak{S}^2_1(r_1, r_2)$ is not essential, and could be removed at the expense of a certain amount of topological complication. Since such complication here would only obscure the main idea, we have decided to omit it.

A similar remark applies to the statement of the theorem, and the following stronger assertion is true: *Of the four possible combinations $\mathfrak{C}_\lambda \cap \mathfrak{C}_\mu \cap \mathfrak{C}_\nu$ for distinct λ, μ, ν, at most two are empty.*

We remark finally that there are three possibilities for the boundary point $\mathfrak{p}$ of $\mathfrak{P}^2$: (i) $\mathfrak{p} = (0, 0)$; (ii) $\mathfrak{p} = (0, r_2)$, $r_2 > 0$; and (iii) $\mathfrak{p} = (r_1, 0)$, $r_1 > 0$. In case (i) the surface $\mathfrak{S}^2_1(r_1, r_2)$ converges to the point $z_1 = z_2 = 0$, in cases (ii) and (iii) to a segment (or a point) of the arc $E[z_k = 0, |z_{3-k}| = r_{3-k}]$. There are thus essentially two cases: convergence of $\mathfrak{S}^2_1(r_1, r_2)$ to a point or to a curve.

5. The proof of the theorem is based on a lemma which we formulate in slightly more general form than we actually require.

Suppose that $\mathfrak{b}^1$ is a (closed) Jordan curve in n-dimensional Euclidean space $\mathfrak{R}^n$ defined by $x_\nu = \chi_\nu(t)$, $0 \le t \le 1$, $\nu = 1, 2, \cdots, n$; and make an arbitrary dissection of $\mathfrak{b}^1$ into four parts $\mathfrak{b}^1_\mu$ such that

$$\mathfrak{b}^1 = \mathop{\mathfrak{S}}_{\mu=1}^{4} \mathfrak{b}^1_\mu, \qquad \mathfrak{b}^1_\mu = E[x_\nu = \chi_\nu(t), t_{\mu-1} \le t \le t_\mu] \tag{5.1}$$
$$(t_{\mu-1} < t_\mu, t_0 = 0, t_4 = 1).$$

We suppose that

$$D(\mathfrak{b}^1_1, \mathfrak{b}^1_3) \ge \alpha_1 > 0, \qquad D(\mathfrak{b}^1_2, \mathfrak{b}^1_4) \ge \alpha_2 > 0, \tag{5.2}$$

where $D[\mathfrak{b}^1_k, \mathfrak{b}^1_{k+2}]$ is the distance between $\mathfrak{b}^1_k$ and $\mathfrak{b}^1_{k+2}$ (that is to say, the minimum distance between any two points $\mathfrak{p}_1$ and $\mathfrak{p}_2$, $\mathfrak{p}_1 \in \mathfrak{b}^1_k$, $\mathfrak{p}_2 \in \mathfrak{b}^1_{k+2}$). Let $\mathfrak{S}^2$ be any continuous surface in $\mathfrak{R}^n$ with boundary $\mathfrak{b}^1$.

LEMMA. *If the area $A(\mathfrak{S}^2)$ of $\mathfrak{S}^2$ exists, then*

$$A(\mathfrak{S}^2) \ge \alpha_1\alpha_2. \tag{5.3}$$

By a well-known theorem [5], [6], there exists a surface

$$\mathfrak{S}_0^2 = E[x_\nu = \chi_\nu(u_1, u_2), 0 < u_k < \alpha_k] \tag{5.4}$$

of *least* area having $\mathfrak{b}^1$ for boundary for which, furthermore, writing

$$g_{jk} = \sum_{\nu=1}^{n} \frac{\partial \chi_\nu}{\partial u_j} \frac{\partial \chi_\nu}{\partial u_k}, \tag{5.5}$$

we have

$$g_{11} = g_{22}, \qquad g_{12} = g_{21} = 0. \tag{5.6}$$

That is to say, $\mathfrak{S}_0^2$ is mapped conformally onto the rectangle $0 < u_k < \alpha_k$.

We write

$$\mathfrak{c}_\mu = \mathfrak{b}_\mu^1 \cap \mathfrak{b}_{\mu+1}^1 \qquad (\mu = 1, 2, 3, 4;\ \mathfrak{b}_5^1 = \mathfrak{b}_1^1). \tag{5.7}$$

By mapping conformally the rectangle $E[0 < u_k < \alpha_k]$ into itself, we may suppose that $\mathfrak{c}_1$, $\mathfrak{c}_2$, $\mathfrak{c}_3$ correspond to $(0, 0)$, $(0, \alpha_1)$, (α_1, α_2) respectively. There are now three possibilities: $\mathfrak{c}_4$ corresponds to the corner $(0, \alpha_2)$, to a point of the line $u_1 = 0$, or to a point of the line $u_2 = \alpha_2$. We may suppose without loss of generality that $\mathfrak{c}_4$ corresponds to a point of $u_1 = 0$ (which may be $(0, \alpha_2)$). By (5.6)

$$A(\mathfrak{S}_0^2) = \int_0^{\alpha_1} \int_0^{\alpha_2} g_{11}\, du_1\, du_2, \tag{5.8}$$

and $L[\chi(\mathfrak{l}_{u_1}^1)]$, the length of the transform by $\chi = [x_\nu = \chi_\nu(u_1, u_2)]$ of the line-segment $\mathfrak{l}_{u_1}^1 = E[u_1 = \text{const.}, 0 < u_2 < \alpha_2]$, is given by

$$L[\chi(\mathfrak{l}_{u_1}^1)] = \int_0^{\alpha_2} g_{11}^{\frac{1}{2}}\, du_2. \tag{5.9}$$

Since $\chi(\mathfrak{l}_{u_1}^1)$ connects $\mathfrak{b}_2^1$ to $\mathfrak{b}_4^1$, we plainly have $L[\chi(\mathfrak{l}_{u_1}^1)] \geq \alpha_2$. Hence, by the inequality of Schwarz,

$$\begin{aligned} \alpha_1^2\alpha_2^2 &\leq \left\{ \int_0^{\alpha_1} \left[\int_0^{\alpha_2} g_{11}^{\frac{1}{2}}\, du_2 \right] du_1 \right\}^2 \\ &\leq \int_0^{\alpha_1}\int_0^{\alpha_2} g_{11}\, du_1\, du_2 \cdot \int_0^{\alpha_1}\int_0^{\alpha_2} du_1\, du_2 = \alpha_1\alpha_2 \cdot A(\mathfrak{S}_0^2) \leq \alpha_1\alpha_2 \cdot A(\mathfrak{S}^2). \end{aligned} \tag{5.10}$$

6. Suppose that the statement of the theorem is false. Then

$$\mathfrak{E}_1 \cap \mathfrak{E}_3 = 0, \qquad \mathfrak{E}_2 \cap \mathfrak{E}_4 = 0. \tag{6.1}$$

By (6.1) there exist two positive numbers ϵ and $r_0 = r_0(\epsilon)$ such that in $\mathfrak{B}^2 = E[(r_1 - p_1)^2 + (r_2 - p_2)^2 \leq r_0^2]$, where (see end of §4) $\mathfrak{p} = (p_1, p_2)$, we have

$$D[\mathfrak{b}_{2,k}^1(r_1, r_2), \mathfrak{b}_{2,k+2}^1(r_1, r_2)] \geq \epsilon, \tag{6.2}$$

where $\mathfrak{b}^1_{2,\nu}(r_1, r_2) = \mathbf{w}\,[\mathfrak{b}^1_{1,\nu}(r_1, r_2)]$. Writing now

$$\mathfrak{S}^2_2(r_1, r_2) = \mathbf{w}\,[\mathfrak{S}^2_1(r_1, r_2)], \tag{6.3}$$

we obtain from the lemma that

$$A\{\mathfrak{S}^2_2(r_1, r_2)\} \geq \epsilon^2 \tag{6.4}$$

for all $(r_1, r_2) \in \mathfrak{P}^2 \cap \mathfrak{B}^2$.

We write

$$f_1 = v_1 + iv_2, \qquad f_2 = v_3 + iv_4, \qquad v_\nu = v_\nu(r_1 e^{i\psi_1}, r_2 e^{i\psi_2}) \qquad (\nu = 1, 2, 3, 4)$$

and

$$g_{jk} = \sum_{\nu=1}^{4} \frac{\partial v_\nu}{\partial \psi_j} \frac{\partial v_\nu}{\partial \psi_k}. \tag{6.5}$$

Then

$$A\{\mathfrak{S}^2_2(r_1, r_2)\} = \iint\limits_{\mathfrak{M}^2(r_1, r_2)} (g_{11}g_{22} - g_{12}^2)^{\frac{1}{2}}\, d\psi_1\, d\psi_2\,. \tag{6.6}$$

Hence, by (6.4) and the inequality of Schwarz, we have, for $(r_1, r_2) \in \mathfrak{P}^2 \cap \mathfrak{B}^2$,

$$\begin{aligned} \epsilon^4 &\leq \left\{ \iint\limits_{\mathfrak{M}^2(r_1, r_2)} (g_{11}g_{22} - g_{12}^2)^{\frac{1}{2}}\, d\psi_1\, d\psi_2 \right\}^2 \\ &\leq \iint\limits_{\mathfrak{M}^2(r_1, r_2)} d\psi_1\, d\psi_2 \cdot \iint\limits_{\mathfrak{M}^2(r_1, r_2)} (g_{11}g_{22} - g_{12}^2)\, d\psi_1\, d\psi_2 \\ &\leq 4\pi^2 \iint\limits_{\mathfrak{M}^2(r_1, r_2)} (g_{11}g_{22} - g_{12}^2)\, d\psi_1\, d\psi_2 \end{aligned} \tag{6.7}$$

since $A[\mathfrak{M}^2(r_1, r_2)] \leq 4\pi^2$. But [3; 476]

$$\begin{aligned} g_{11}g_{22} - g_{12}^2 &= \left| \frac{\partial(f_1, f_2)}{\partial(z_1, z_2)} \right|^2 (r_1 r_2)^2 + \left[Im\left(\sum_{k=1}^{2} \frac{\partial f_k}{\partial z_1} \frac{\partial f_k^*}{\partial z_2^*} e^{i(\psi_1 - \psi_2)} \right) \right]^2 (r_1 r_2)^2 \\ &\leq (r_1 r_2)^2 [J(z_1, z_2) + H(z_1, z_2)] \end{aligned} \tag{6.8}$$

(* denotes the conjugate) and so by (6.7)

$$\iint\limits_{\mathfrak{M}^2(r_1, r_2)} (J + H) r_1\, d\psi_1 r_2\, d\psi_2 \geq \frac{\epsilon^4}{4\pi^2} \cdot \frac{1}{r_1 r_2} \qquad ((r_1, r_2) \in \mathfrak{P}^2 \cap \mathfrak{B}^2). \tag{6.9}$$

Finally, integrating (6.9) with respect to r_1, r_2 over $\mathfrak{P}^2 \cap \mathfrak{B}^2$, we find

$$\begin{aligned}(6.10)\qquad &\iiiint\limits_{\mathfrak{R}} (J + H)\, d\omega \\ &\geq \iint\limits_{\mathfrak{P}^2 \cap \mathfrak{B}^2} \left\{ \iint\limits_{\mathfrak{M}^2(r_1, r_2)} (J + H)\, d\psi_1\, d\psi_2 \right\} r_1\, dr_1\, r_2\, dr_2 \geq \frac{\epsilon^4}{4\pi^2} \iint\limits_{\mathfrak{P}^2 \cap \mathfrak{B}^2} \frac{dr_1 dr_2}{r_1 r_2} = \infty\end{aligned}$$

by (4.1), a result which contradicts (4.7). This completes the proof of the theorem.

7. We observe that the kernel of the method lies in the inequalities (5.10) and (6.7). For conformal mapping there is the relation (5.10), and for quasi-conformal mapping a similar inequality. For pseudo-conformal mapping in the space of two complex variables, there exist two analogous inequalities: one connects "B-area" of surfaces and 4-dimensional volume, and the other area of surfaces in the ordinary sense and the functional (4.7). (See [4; 147–149] and [2], [3].) The problem arises to characterize more general classes of mappings with this property.

Bibliography

1. Stefan Bergman, *Ueber eine Abschätzung von meromorphen Funktionen zweier komplexen Veränderlichen in Bereichen mit ausgezeichneter Randfläche*, Travaux de l'Institut Mathématique de Tbilissi, vol. 1(1937), pp. 187–204.
2. Stefan Bergman, *Conformal mapping of multiply-connected domains* (Partial differential equations, advanced topics), Publication of Brown University, 1941.
3. Stefan Bergman, *Sur quelques propriétés des transformations par un couple des fonctions de deux variables complexes*, Rend. R. Accad. Naz. Lincei, (6), vol. 19(1934), pp. 474–478.
4. Stefan Bergman, *Theory of pseudo-conformal transformations and its connection with differential geometry*, Notes of lectures delivered at the Massachusetts Institute of Technology, 1939–1940 (Brown University Library).
5. J. Douglas, *Solution of the problem of Plateau*, Transactions of the American Mathematical Society, vol. 33(1931), pp. 263–321.
6. T. Radó, *On the problem of Plateau*, Ergebnisse der Mathematik, Berlin, vol. 2(1933).
7. D. C. Spencer, *On finitely mean valent functions*. II, Transactions of the American Mathematical Society, vol. 48(1940), pp. 418-435.
8. W. F. Whitmore, *Convergence theorems for functions of two complex variables*, American Journal of Mathematics, vol. 62(1940), pp. 687–696.

Brown University and Stanford University.

CONTINUED FRACTIONS WITH COMPLEX ELEMENTS

By Walter Leighton and W. J. Thron

Introduction. In this paper are obtained new conditions which are sufficient for the convergence of a continued fraction of the form (1.1). In particular, we determine a new set of *convergence regions* for (1.1), that is, regions R of the complex plane with the following property: a sufficient condition for the convergence of (1.1) is that the set $a_1, a_2, \cdots$ belong to R. The results established here are improvements on certain earlier results reported to the American Mathematical Society May 9, 1941 (cf. Bulletin of that Society, vol. 47(1941), p. 557). The methods of the paper include what is believed to be the first application of the theory of normal families to the study of convergence of complex continued fractions.

In §1 and §2 certain preliminary basic results are obtained. The main results of the paper appear in Theorems 3.1 and 3.2 in §3. Theorem 3.1 provides conditions on the numbers a_n and the variable z which insure the convergence of $K(a_n z/1)$ to an analytic function of z. As an immediate consequence of this there appears in Theorem 3.2 a new family of parabolic convergence regions. In §4 the question of "bestness" is discussed. In §5 results on value regions for continued fractions are presented.

1. **A fundamental lemma.** If the complex numbers $a_1, a_2, \cdots$ lie in a region E in the complex plane, E will be said to be an *element region* associated with the continued fraction

$$1 + \frac{a_1}{1 +} \frac{a_2}{1 +} \cdots . \tag{1.1}$$

If V is a region in which the values of all the approximants

$$\frac{A_n}{B_n} = 1 + \frac{a_1}{1 +} \frac{a_2}{1 +} \cdots \frac{a_n}{1} \qquad (n = 1, 2, \cdots)$$

lie, V will be said to be a *value region* corresponding to the element region E. It is not assumed that the region V contains only values taken on by the approximants A_n/B_n. Further, no assumption is made concerning the convergence of the continued fraction (1.1).

A sufficient condition that a set V be the value region corresponding to an element region is given in the following lemma, due to Scott and Wall [3].

Received March 10, 1942. Presented to the American Mathematical Society on September 2, 1941 and in revised form on April 17, 1942.

LEMMA 1.1. *If two sets of complex numbers E and V are given, such that*

(i) $$1 + a \in V, \text{ if } a \in E,$$

(ii) $$1 + a/v \in V, \text{ if } a \in E \text{ and } v \in V,$$

then V is a value region corresponding to the element region E

A brief proof may be given by induction. It follows from (i) that every number $1 + a_1/1$ belongs to V. Assume now that all numbers $1 + a_1/1 + a_2/1 + \cdots + a_k/1$ belong to V for any choice of the k numbers $a_1, a_2, \cdots, a_k$ $(k = 1, 2, \cdots, m)$ from E. It then follows from (ii) that all numbers $1 + a_1/1 + a_2/1 + \cdots + a_{m+1}/1$ belong to V for any choice of the $m + 1$ numbers $a_1, \cdots, a_{m+1}$.

The following corollary is immediate.

COROLLARY 1.1. *If two sets E and V are given, such that*

(i)′ $$1 \in V,$$

(ii) $$1 + a/v \in V, \text{ if } a \in E \text{ and } v \in V,$$

then V is a value region corresponding to the element region E.

2. **Construction of element regions corresponding to a given value region.** The problem of constructing a precise value region corresponding to a given element region does not appear in general to be simple. In this section we shall describe, however, a general method for determining an element region corresponding to a given value region.

Let V be a set of complex numbers whose elements we shall denote by v. Let b be an arbitrary complex number. The set $V + b$, we shall understand, consists of all numbers $v + b$, where v belongs to V. Similarly, the set $b \cdot V$ shall consist of all numbers $b \cdot v$, where v is any member of V. Finally, let B be a set of complex numbers whose elements are denoted by b. The set $D[b \cdot V]$ denotes the point-set intersection of all sets $b \cdot V$, where b is an element of B.

Using this notation, we have the following theorem.

THEOREM 2.1. *Let V be a region in the complex plane, such that $1 \in V$. If the set $E = D[v(V - 1)]$ is not empty, V is a value region corresponding to the element region E.*

This theorem is an immediate consequence of Corollary 1.1. Since $1 \in V$, (i)′ is satisfied. From the definition of the set E it follows that for any given $a \in E$ and $v \in V$ an element $v' \in V$ can be found such that $a = v'(v - 1)$. Then $v' = 1 + a/v$, and $v' \in V$ for any choice of a and v; hence (ii) is satisfied and the proof is complete.

We shall now be concerned with the actual construction of regions E arising from certain simple V regions.

Let V be the convex region bounded by two rays emanating from $z = d$, where d is a positive number satisfying $\frac{1}{2} \leq d < 1$. The angle θ between the two rays is to be $\leq \pi$. Further, it is assumed that the two rays do not lie both above or both below the real axis and that the points on the real axis with values which $\geq d$ are contained in the region V. Denote the rays by a and b. By a' and b', respectively, we denote the images of the rays a and b under the transformation $w = z - 1$. For these regions the following lemmas hold.

LEMMA 2.1. *The region $D[v(V - 1)]$, where v ranges over the whole set V, is equal to the region $D[v'(V - 1)]$, where v' ranges over the boundary of V only.*

It suffices to show that for any set $v(V - 1)$ a set $v'(V - 1)$ contained in $v(V - 1)$ can be found, where v' is a point of the boundary of V. It is clear that such a number v' is obtained by taking the number lying on the boundary and satisfying the condition $\arg(v') = \arg(v)$. These two numbers will satisfy the inequality $|v'| \leq |v|$.

The set $(V - 1)$ is subjected to the same rotation by multiplication by v as by multiplication by v'. Further, as $v = 0$ is contained in the set $(V - 1)e^{i\theta}$, the stretching effected by $|v| \geq |v'|$ insures that the set $v(V - 1)$ contains the set $v'(V - 1)$.

LEMMA 2.2. *If we denote by $z = x + iy$ the points on a straight line*

$$x \cos \theta + y \sin \theta - p = 0,$$

the points $z \cdot re^{i\varphi} = u + iv$ lie on the straight line

$$u \cos(\theta + \varphi) + v \sin(\theta + \varphi) - rp = 0.$$

It is well known that, under a rotation as well as under a stretching, a straight line is transformed into a straight line. Further, multiplication by a complex number gives rise to a rotation followed by a stretching, each with respect to the origin. The truth of the lemma is now evident.

LEMMA 2.3. *The boundary of the region $D[v(V - 1)]$ is formed by the two families of straight lines*:

(1) *line a' multiplied by all points on ray b,*
(2) *line b' multiplied by all points on ray a.*

If a letter, originally assigned to a ray, is used for a line, it is assumed to mean the line of which the ray is a part. Further, let us denote by a'' and b'' the images, respectively, of the rays a and b under the transformation $w = d(z - 1)$, where d is the number previously defined. Finally, denote by s a ray obtained by multiplying the ray a' by a number representing a point on the ray a. With these preparations we are ready to prove the lemma.

It will be shown that no ray s has a point in common with the interior of the set $d \cdot (V - 1)$. This set contains the set $D[v(V - 1)]$. In a similar way it can

be shown that no ray of the family obtained by multiplying ray b' by numbers representing points on the ray b has any point in common with the interior of the set $d\cdot(V-1)$.

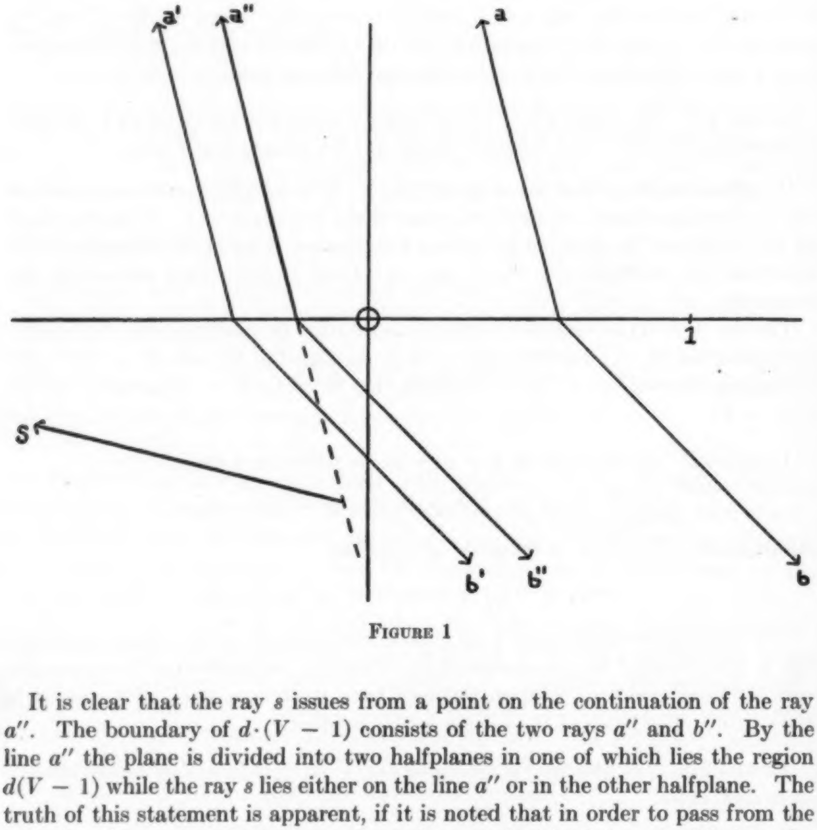

FIGURE 1

It is clear that the ray s issues from a point on the continuation of the ray a''. The boundary of $d\cdot(V-1)$ consists of the two rays a'' and b''. By the line a'' the plane is divided into two halfplanes in one of which lies the region $d(V-1)$ while the ray s lies either on the line a'' or in the other halfplane. The truth of this statement is apparent, if it is noted that in order to pass from the ray a'' to the ray s one must move through an angle θ $(0 \leq \theta < \pi)$ while in order to pass from the ray a'' to the ray b'' one must move through an angle φ $(-\pi \leq \varphi < 0)$. The proof of the lemma is complete.

Let the region V be bounded by the rays

$$(2.1)\qquad \begin{aligned} &\text{(a)} \quad x\cos\beta + y\sin\beta - d\cos\beta = 0, \qquad y \geq 0, \\ &\text{(b)} \quad x\cos\gamma + y\sin\gamma - d\cos\gamma = 0, \qquad y \leq 0. \end{aligned}$$

Here $-\frac{1}{2}\pi < \beta < \frac{1}{2}\pi$, $-\frac{1}{2}\pi < \gamma < \frac{1}{2}\pi$, $\beta - \gamma \leq 0$, d is real and $\frac{1}{2} \leq d < 1$. The lemmas just proved apply to this region.

To obtain the boundary of the region $E = D[v(V-1)]$ one has to consider the two families of lines obtained by multiplying the line

$$x \cos (\pi + \beta) + y \sin (\pi + \beta) - (1 - d) \cos \beta = 0$$

by points

$$z = \frac{d \cos \gamma}{\cos (\theta - \gamma)} e^{i\theta}, \qquad 0 > \theta > -\tfrac{1}{2}\pi + \gamma$$

and the line

$$x \cos (\pi + \gamma) + y \sin (\pi + \gamma) - (1 - d) \cos \gamma = 0$$

by points

$$z = \frac{d \cos \beta}{\cos (\theta - \beta)} e^{i\theta}, \qquad 0 < \theta < \tfrac{1}{2}\pi + \beta.$$

After a simple change in the parameter the resulting families are seen to be members of the family

$$x \cos (\varphi + \beta + \gamma) + \sin (\varphi + \beta + \gamma) + \frac{d(1 - d) \cos \beta \cos \gamma}{\cos \varphi} = 0, \tag{2.2}$$

where $-\frac{1}{2}\pi < \varphi < \frac{1}{2}\pi$.

Differentiation with respect to φ leads to

$$\begin{aligned} -x \sin (\varphi + \beta + \gamma) &+ y \cos (\varphi + \beta + \gamma) \\ &+ \frac{d(1 - d) \cos \beta \cos \gamma \sin \varphi}{\cos^2 \varphi} = 0. \end{aligned} \tag{2.3}$$

Relations (2.2) and (2.3) are equivalent to

$$x = -\frac{d(1 - d) \cos \beta \cos \gamma}{\cos^2 \varphi} \cos (2\varphi + \beta + \gamma),$$

$$y = -\frac{d(1 - d) \cos \beta \cos \gamma}{\cos^2 \varphi} \sin (2\varphi + \beta + \gamma).$$

Now set $x + iy = re^{i\theta}$; then we have

$$r = (x^2 + y^2)^{\frac{1}{2}} = \frac{d(1 - d) \cos \beta \cos \gamma}{\cos^2 \varphi},$$

and

$$\cos \theta = -\cos (2\varphi + \beta + \gamma),$$

$$\sin \theta = -\sin (2\varphi + \beta + \gamma);$$

hence $\theta + \pi = 2\varphi + \beta + \gamma$ and $\varphi = \frac{1}{2}(\theta - (\beta + \gamma)) + \frac{1}{2}\pi$. From this it follows that

$$\cos^2\varphi = \sin^2 \tfrac{1}{2}(\theta - (\beta + \gamma)) = \tfrac{1}{2}(1 - \cos (\theta - (\beta + \gamma))).$$

It follows that E contains the interior of the parabola

$$r = \frac{2d(1 - d) \cos \beta \cos \gamma}{1 - \cos (\theta - (\beta + \gamma))}. \tag{2.4}$$

For a fixed angle $\beta + \gamma$ one obtains the largest possible parabola by setting $d = \frac{1}{2}$, $\beta = \gamma$. Then V is the halfplane containing the points $v = Re^{i\varphi}$ satisfying $R \geq \frac{1}{2} \cos \beta/\cos(\varphi - \beta)$. We can now state the following result.

THEOREM 2.2. *The value region $V(\beta)$ corresponding to the parabolic element region $E(\beta)$, where $re^{i\theta} \varepsilon E(\beta)$ if*

$$r \leq \frac{\frac{1}{2} \cos^2 \beta}{1 - \cos (\theta - 2\beta)}, \tag{2.5}$$

is the halfplane defined by the relation $Re^{i\varphi} \varepsilon V(\beta)$ if

$$R \geq \frac{\cos \beta}{2 \cos (\varphi - \beta)}, \qquad -\tfrac{1}{2}\pi + \beta < \varphi < \tfrac{1}{2}\pi + \beta. \tag{2.6}$$

It is of interest to note that all the parabolas (2.5) have the point $z = -\frac{1}{4}$ as a point on the boundary.

The following lemma will be useful.

LEMMA 2.4. *The value region V corresponding to the element region E, consisting of elements a satisfying $a \varepsilon E(\beta)$, $| a | < M$, is that part of the previously defined halfplane $V(\beta)$ for which $| v | < 1 + 2M/\cos \beta$.*

For the proof, note first that $| 1 + a | < M + 1 < \dfrac{2M}{\cos \beta} + 1$. Further, as v lies in the halfplane $R \geq \frac{1}{2} \cos \beta \sec (\varphi - \beta)$, we have $| v | > \frac{1}{2} \cos \beta$. For the second and higher approximants we then have $| v_n - 1 | \cdot | v_{n-1} | = | a |$; hence $| v_n - 1 | < 2M \sec \beta$ and $| v_n | < 2M \sec \beta + 1$.

3. **Application of the theory of normal families.** The foundations have now been laid for applying the theory of normal families to continued fractions. The functions to be considered are the approximants $A_n(z)/B_n(z)$ of the continued fraction

$$1 + \frac{a_1 z}{1 +} \frac{a_2 z}{1 +} \frac{a_3 z}{1 +} \cdots, \tag{3.1}$$

where the numbers a_n lie in a closed bounded region in the interior of a parabola $E(\beta)$ and z is a complex variable.

Let $|a_n| < M$ $(n \geq 1)$ and let $k > 1$ be a real number such that $a_n k \varepsilon E(\beta)$ for all n. Such a number k evidently exists. One could set $k = 1 + \delta/M$, where δ is the greatest lower bound of the distances of the numbers a_n from the boundary of $E(\beta)$.

Since the parabola $E(\beta)$ is convex and contains the point $z = 0$ in its interior, $a_n r \varepsilon E(\beta)$, if $a_n \varepsilon E(\beta)$ and r is real, positive, and $< k$.

Further, multiplication by

$$\frac{\cos^2 (\beta + \frac{1}{2}\theta)}{\cos^2 \beta} e^{i\theta}$$

transforms the parabola $E(\beta)$ into $E(\beta + \frac{1}{2}\theta)$. These remarks make it evident that if the variable z lies in the region D, where

$$z = re^{i\theta} \varepsilon D \text{ if } r < \frac{k \cos^2 (\beta + \frac{1}{2}\theta)}{\cos^2 \beta},$$

the elements $a_n z$ of the continued fraction (3.1) lie in a closed bounded region in the interior of one of the parabolas $E(\beta)$.

A consequence of Lemma 2.4 is that the functions $A_n(z)/B_n(z)$ are finite for all values of z in D, and since the approximants are rational functions of z, it follows that they are analytic in D.

The values on the real axis $< \frac{1}{2}$ are not assumed by any function of the sequence $A_n(z)/B_n(z)$ for any value $z \varepsilon D$. This follows from Theorem 2.2. The sequence of approximants, therefore, is a normal family of analytic functions in D.

Further, for $|z| < 1/4M$, all elements of the continued fraction (3.1) satisfy the condition $|a_n z| < \frac{1}{4}$ $(n \geq 1)$, and hence according to a theorem due to Worpitzky [4] the continued fraction, that is, its sequence of approximants, converges to a finite value.

The intersection of the circle $|z| < 1/4M$ with the region D contains an open set. Montel's [1] generalization of the Stieltjes-Vitali theorem can therefore be applied and leads to the conclusion that the sequence of approximants $A_n(z)/B_n(z)$ converges uniformly (and hence to an analytic function) for the whole region D. We have now proved the following theorem.

THEOREM 3.1. *If the numbers $a_n = re^{i\theta}$ lie in a closed bounded region in the interior of one of the parabolas $E(\beta)$ $(-\frac{1}{2}\pi < \beta < \frac{1}{2}\pi)$,*

$$r \leq \frac{\frac{1}{2} \cos^2 \beta}{1 - \cos (\theta - 2\beta)},$$

then for the values of the variable $z = Re^{i\varphi}$ that satisfy the relation

$$R < \frac{k \cos^2 (\beta + \frac{1}{2}\varphi)}{\cos^2 \beta}$$

the continued fraction

$$1 + \frac{a_1 z}{1 +} \frac{a_2 z}{1 +} \cdots$$

converges uniformly to an analytic function.

Recalling that $z = 1$ is in D, we have the following theorem which is an immediate consequence of the preceding.

THEOREM 3.2. *If the elements $a_n = re^{i\theta}$ lie in a closed bounded region in the interior of the parabolic region $E(\beta)$ $(-\frac{1}{2}\pi < \beta < \frac{1}{2}\pi)$,*

$$r \leq \frac{\frac{1}{2}\cos^2\beta}{1 - \cos(\theta - 2\beta)},$$

the continued fraction

$$1 + \frac{a_1}{1 +} \frac{a_2}{1 +} \cdots$$

converges.

4. **Bestness.** In this section we shall consider two closely related examples. The first will show that the halfplane $V(\beta)$ is actually filled by values assumed by continued fractions (1.1) with elements in the parabola $E(\beta)$. The second example shows that in a certain sense the parabolas $E(\beta)$ are the best convergence regions obtainable.

To this end consider the periodic continued fraction

$$1 + \frac{a_1}{1 +} \frac{a_2}{1 +} \frac{a_1}{1 +} \cdots, \tag{4.1}$$

where

$$a_1 = \tfrac{1}{2}e^{i(\frac{1}{2}\pi+\beta)}[\sin\beta + (1 + 2d)] + e^{i2\beta}(d + d^2 + k),$$

$$a_2 = \tfrac{1}{2}e^{i(\frac{1}{2}\pi+\beta)}[\sin\beta - (1 + 2d)] + e^{i2\beta}(d + d^2 + k).$$

It is easily verified that if $d \geq -\frac{1}{2}$ and $k > 0$, a_1 and a_2 lie in the parabola. For $k = 0$ they lie on the boundary and the line passing from a_1 to a_2 makes an angle $\frac{1}{2}\pi + \beta$ with the axis of the parabola.

From the periodic theory [2] it is known that the value of the continued fraction (4.1) is

$$v = \tfrac{1}{2}\{1 + a_1 - a_2 \pm [(1 + a_1 - a_2)^2 + 4a_2]^{\frac{1}{2}}\}.$$

The substitution of the values of a_1 and a_2 leads to

$$v = \tfrac{1}{2}(1 + (1 + 2d)e^{i(\frac{1}{2}\pi+\beta)} \pm 2k^{\frac{1}{2}}e^{i\beta}).$$

The correct sign for the last term is the plus sign ($v \in V(\beta)$). It is then easily seen that with the permissible range of values for k and d every value of the halfplane $V(\beta)$ is assumed by a continued fraction (4.1).

Let us now consider a similar continued fraction

$$1 + \frac{b_1}{1 +} \frac{b_2}{1 +} \frac{b_1}{1 +} \cdots , \tag{4.2}$$

where

$$b_1 = \tfrac{1}{2}e^{i(\frac{1}{2}\pi+\beta)} [\sin \beta + (1 + 2d + \epsilon)] + (d + d^2)e^{i2\beta},$$

$$b_2 = \tfrac{1}{2}e^{i(\frac{1}{2}\pi+\beta)} [\sin \beta - (1 + 2d + \epsilon)] + (d + d^2)e^{i2\beta}.$$

Here ϵ is supposed to be a complex number. For certain values of ϵ the continued fraction (4.2) may not converge.

It is known [2; 276] that the continued fraction (4.2) diverges if the following two conditions are satisfied:

(a) $[(1 + b_1 - b_2)^2 + 4b_2]^{\frac{1}{2}} \neq 0,$

(b) $$| 1 + b_1 + b_2 + [(1 + b_1 - b_2)^2 + 4b_2]^{\frac{1}{2}} | = | 1 + b_1 + b_2 - [(1 + b_1 - b_2)^2 + 4b_2]^{\frac{1}{2}} |.$$

Condition (b) is satisfied if the number $1 + b_1 + b_2$ regarded as a vector is perpendicular to the vector represented by $[(1 + b_1 - b_2)^2 + 4b_2]^{\frac{1}{2}}$.

We have

$$1 + b_1 + b_2 = \sin \beta\, (1 + 2d)^2 e^{i(\frac{1}{2}\pi+\beta)} + 1 + 2(d + d^2);$$

this expression does not involve ϵ. Further,

$$(1 + b_1 - b_2)^2 + 4b_2 = e^{i(\pi+2\beta)} [2\epsilon(1 + 2d) + \epsilon^2].$$

In every neighborhood of the number zero we can find an ϵ satisfying both conditions (a) and (b), and hence the continued fraction (4.2) diverges.

It is therefore impossible to find a region different from $E(\beta)$ containing a parabola $E(\beta)$ in its interior and having with respect to convergence the same properties as $E(\beta)$. It is further impossible to find a convergence region, containing in its interior neighborhoods of both the points b_1 and b_2.

5. **Value regions.** The work of the previous sections gives some immediate results concerning the values of the continued fraction (1.1).

The value of a continued fraction is the limit of the value of its approximants if this limit exists. Let E be an element region and V a corresponding value region. It is then clear that the value of a convergent continued fraction (1.1) with elements in E must lie in the closure of V.

THEOREM 5.1. *If the elements $a_n = re^{i\theta}$ of the continued fraction* (1.1) *lie in a closed bounded region in the interior of the parabolic region* (2.5), *the continued fraction converges to a value $v = Re^{i\varphi}$ in the halfplane* (2.6). *Every value in this halfplane is taken by at least one continued fraction* (1.1) *with elements in the region* (2.5).

The first part of the theorem is a consequence of Theorem 2.2 and the last part follows from the first example in §4.

A consequence of Lemma 2.4 is the following result.

THEOREM 5.2. *If the elements a_n of the continued fraction* (1.1) *satisfy in addition to the conditions of Theorem* 5.1 *the condition $|a_n| < M$ $(n \geq 1)$, then the continued fraction converges to a value v, satisfying the condition $|v| < 1 + 2M/\cos\beta$.*

Finally, it is of interest to consider the value region corresponding to the element region defined by

$$r \leq \frac{k}{1 - \cos(\theta - 2\alpha)}, \qquad 0 < k < \tfrac{1}{2}\cos^2\alpha.$$

From the discussion in §2, it follows that the corresponding value region must be the part common to all regions $V(\beta, d)$ defined by the two relations $r \geq d/\cos(\beta - \theta)$ and $r \geq d/\cos(2\alpha - \beta - \theta)$, where β and d, according to (2.4), must satisfy the relation

$$2d(1 - d)\cos\beta\cos(2\alpha - \beta) = k.$$

BIBLIOGRAPHY

1. P. A. MONTEL, *Leçons sur les familles normales de fonctions analytiques et leurs applications*, Paris, 1927.
2. O. PERRON, *Die Lehre von den Kettenbrüchen*, Leipzig, 1929.
3. W. T. SCOTT AND H. S. WALL, *Value regions for continued fractions*, Bulletin of the American Mathematical Society, vol. 47(1941), pp. 580–585.
4. WORPITZKY, Jahresbericht Friedrichsgymnasium und Realschule, Berlin, 1865, pp. 3–39.

THE RICE INSTITUTE.

EFFECTIVE PARAMETERS

By Arthur B. Brown

1. **Introduction.** Under the classical treatment of essential parameters [1], [3] for a set of functions $f_1(x_1, \cdots, x_n, \alpha_1, \cdots, \alpha_m), f_2(x, \alpha), \cdots, f_r(x, \alpha)$, the parameters $\alpha_1, \alpha_2, \cdots, \alpha_m$ are called *essential* if there do not exist $m - 1$ functions of them, say $A_1(\alpha), \cdots, A_{m-1}(\alpha)$, and r functions $F_1, \cdots, F_r$ such that

$$(1.1) \qquad f_i(x; \alpha_1, \cdots, \alpha_m) \underset{(x,\alpha)}{\equiv} F_i[x; A_1(\alpha), \cdots, A_{m-1}(\alpha)] \qquad (i = 1, \cdots, r).$$

The treatment includes an algorithm for determining the number of essential parameters in terms of which the given functions can be expressed. If that number is $m - 1$, for example, then the final set of functions are the F_i in (1.1).

A weakness in this treatment is that the functions F_i are in general unknown. (The fact that the A's are in general unknown functions of the α's seems of little consequence.) In this paper we show that the same algorithm leads to like results, where in place of identities of the form (1.1) we have similar identities in the x's only. But in our case *the F's are known functions*; in fact, $F_i = f_i$ with $\alpha_j = A_j$ for a known set of subscripts j, with the remaining α's constant (Theorem 8.1). All r-tuples $f_1, f_2, \cdots, f_r$ of functions of the x's are obtained in this way which can be obtained by varying all the α's, and, incidentally, without duplication.

In all the work certain "singular" points must be avoided, both in the classical and in the present treatment. A discussion of the singular points is included.

A second deficiency in the classical treatment is the failure to establish a minimum number of parameters in terms of which the given r-tuple of families of functions of $(x_1, \cdots, x_n)$ can be expressed by means of differentiable functions; for this is done only under the restriction that the new parameters be functions of the old. In this paper we show without restriction, except as to the class of the given functions and the singularity of the points, that no smaller number of parameters can be sufficient (Theorem 10.5).

2. **Preliminaries.** Let functions $f_1(x_1, \cdots, x_n, \alpha_1, \cdots, \alpha_m) \equiv f_1(x, \alpha)$, $f_2(x, \alpha), \cdots, f_r(x, \alpha)$ be given, real and of class C^{N+3}, in a neighborhood of $(x_1^0, \cdots, x_n^0, \alpha_1^0, \cdots, \alpha_m^0)$; N is to be specified later; $n, m, r \geq 1$. If the functions are analytic, no restriction to real functions is necessary. We consider the f's as functions of $(x_1, \cdots, x_n)$, with $\alpha_1, \cdots, \alpha_m$ as parameters. Let us denote by $\{f(x, \alpha)\}$ the r-tuple $f_1(x, \alpha), \cdots, f_r(x, \alpha)$ of functions. In the case of a set containing only one function, we omit the braces.

Received March 31, 1942. Presented to the American Mathematical Society, October 26, 1940, under the title *On the number of independent parameters*, and October 25, 1941, under the title *Independent parameters for sets of functions*.

DEFINITION. The arguments $\alpha_1, \cdots, \alpha_k$, $k \leq m$ (or any other subset of the α's if they satisfy corresponding conditions) will be called an *effective* set of parameters for $\{f(x, \alpha)\}$ at (x^0, α^0) if, for every $(\alpha_1^1, \cdots, \alpha_k^1)$, $(\alpha_1^2, \cdots, \alpha_k^2)$ in some neighborhood of $(\alpha_1^0, \cdots, \alpha_k^0)$, with $(\alpha_1^1, \cdots, \alpha_k^1) \neq (\alpha_1^2, \cdots, \alpha_k^2)$, not all the identities

$$f_i(x; \alpha_1^1, \cdots, \alpha_k^1, \alpha_{k+1}^0, \cdots, \alpha_m^0) \underset{(x)}{\equiv} f_i(x; \alpha_1^2, \cdots, \alpha_k^2, \alpha_{k+1}^0, \cdots, \alpha_m^0)$$

are valid for (x) in any neighborhood of (x^0).

The arguments $\alpha_1, \cdots, \alpha_k$ will be called a *sufficient* set of parameters for $\{f(x, \alpha)\}$ at (x^0, α^0) if, for (x) in some neighborhood of (x^0), the r-tuples $\{f(x; \alpha_1, \cdots, \alpha_k, \alpha_{k+1}^0, \cdots, \alpha_m^0)\}$ of functions of (x) with $(\alpha_1, \cdots, \alpha_k)$ in any given neighborhood of $(\alpha_1^0, \cdots, \alpha_k^0)$ include all r-tuples $\{f(x; \alpha_1, \cdots, \alpha_m)\}$ of functions of (x) with $(\alpha_1, \cdots, \alpha_m)$ in a corresponding neighborhood of $(\alpha_1^0, \cdots, \alpha_m^0)$.

A set of parameters will be called *complete* for $\{f(x, \alpha)\}$ at (x^0, α^0) if it is both effective and sufficient for $\{f(x, \alpha)\}$ at (x^0, α^0).

Remark. It follows fairly easily from the first definition that, if all m parameters form an effective set, then they are essential. We shall see later that the number of parameters in a complete set in general equals the minimum number of essential parameters in terms of which the given functions can be expressed. (Cf. the remarks just before Theorem 5.9.)

The following theorems are obvious.

THEOREM 2.1. *If $\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_k}$ are effective for any one function f_i of $\{f\}$ at (x^0, α^0), then they are effective for $\{f\}$ at (x^0, α^0).*

THEOREM 2.2. *If $\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_k}$ are sufficient for $\{f\}$ at (x^0, α^0), then they are sufficient for each individual function of $\{f\}$ at (x^0, α^0).*

While throughout the paper the wording will be for the case of real functions and real variables, all the work goes through for the case of complex variables if the f's are analytic, with no change in proof.

3. **The matrix** (M).

In this section we give an algorithm for the construction of a matrix (M), whose rank turns out to equal the number of parameters in a complete set. Later we show that this matrix can be replaced by the matrix of the classical theory.

All derivatives of a set of r dependent variables, say $z_1, z_2, \cdots, z_r$, with respect to the x's, are first ordered, as follows. We use the symbol δ to denote differentiations with respect to the x's, and begin by ordering the δ's. If

$$\delta_a = \frac{\partial^{i_1 + \cdots + i_n}}{\partial x_1^{i_1} \cdots \partial x_n^{i_n}}, \qquad \delta_b = \frac{\partial^{j_1 + \cdots + j_n}}{\partial x_1^{j_1} \cdots \partial x_n^{j_n}},$$

with $i_1 + \cdots + i_n < j_1 + \cdots + j_n$, then δ_a precedes δ_b. If $i_1 + \cdots + i_n = j_1 + \cdots + j_n$ and $i_1 = j_1$, $i_2 = j_2$, $\cdots$, $i_{k-1} = j_{k-1}$, but $i_k < j_k$, again δ_a precedes δ_b.

We denote by δ_1 the identity operator; in other words, $\delta_1 z_s = z_s$. If now the operators in order are δ_1, δ_2, δ_3, $\cdots$, then the derivatives in order are $\delta_1 z_1$, $\delta_1 z_2$, $\cdots$, $\delta_1 z_r$; $\delta_2 z_1$, $\cdots$, $\delta_2 z_r$; $\delta_3 z_1$, $\cdots$, $\delta_3 z_r$; $\cdots$. If a derivative u precedes a derivative v in the ordering, then u is said to be lower than v, and v to be higher than u. We use some of the notations and ideas of the Riquier theory for partial differential equations, but do not actually use the theory. Cf. [4].

If u denotes $\delta_i z_s$, then δu denotes $\delta(\delta_i z_s)$. It is easily seen that if u is higher than v, then for any δ, δu is higher than δv.

Later, where convenient, we shall use the symbol D to denote differentiations with respect to the x's.

The first "trial" row of (M) is

$$\frac{\partial f_1}{\partial \alpha_1} \quad \frac{\partial f_1}{\partial \alpha_2} \quad \cdots \quad \frac{\partial f_1}{\partial \alpha_m}.$$

If these derivatives are all identically zero in x, α in the neighborhood of (x^0, α^0), we remove the row, and say that z_1 is *rejected* (rejected derivative of z_1 of zeroth order). If the derivatives are not all identically zero in x, α in the neighborhood of (x^0, α^0), but all equal zero at (x^0, α^0), then we shall call (x^0, α^0) a *singular point*. We assume that (x^0, α^0) is not a singular point. As more rows are introduced, we shall have other ways in which (x^0, α^0) may be a singular point. In all cases we shall assume that the point is non-singular. (The locus of singular points in (x, α)-space will be discussed in §9.) If, finally, $\frac{\partial f_1}{\partial \alpha_1}$, $\cdots$, $\frac{\partial f_1}{\partial \alpha_m}$ are not all zero at (x^0, α^0), the row is retained in the matrix, and we say that z_1 is *accepted*.

Now z_2 (undifferentiated) is the next higher derivative. We consider

$$\frac{\partial f_2}{\partial \alpha_1} \quad \frac{\partial f_2}{\partial \alpha_2} \quad \cdots \quad \frac{\partial f_2}{\partial \alpha_m}$$

as next trial row for the matrix. However, instead of discussing this row in detail, let us pass on to the general situation. Suppose that, after testing the rows of derivatives with respect to α_1, $\cdots$, α_m of the functions $\delta_1 f_1 \equiv f_1$, $\delta_1 f_2$, $\cdots$, $\delta_1 f_r$; $\delta_2 f_1$, $\delta_2 f_2$, $\cdots$, $\delta_2 f_r$; $\cdots$, $\delta_s f_t$, some of the derivatives $\delta_1 z_1 = z_1$, $\delta_1 z_2$, $\cdots$, $\delta_1 z_r$; $\delta_2 z_1$, $\delta_2 z_2$, $\cdots$, $\delta_2 z_r$; $\cdots$, $\delta_s z_t$ have been accepted, the rest rejected. We now test the row

$$\frac{\partial}{\partial \alpha_1}(\delta_c f_d) \quad \frac{\partial}{\partial \alpha_2}(\delta_c f_d) \quad \cdots \quad \frac{\partial}{\partial \alpha_m}(\delta_c f_d),$$

where $c = s$ and $d = t + 1$ if $t \neq r$; $c = s + 1$ and $d = 1$ if $t = r$. If at (x^0, α^0) this trial row increases the rank of the matrix, $\delta_c z_d$ is *accepted* and the row becomes

permanently part of the matrix. If the row increases the rank for some points in every neighborhood of (x^0, α^0), but not at (x^0, α^0), then (x^0, α^0) is a *singular point*, and, as above, we assume that that is not the case. If, finally, the row does not increase the rank at any point in a neighborhood of (x^0, α^0), we say that $\delta_c z_d$ is rejected, and the row is removed from the matrix. *Thus at each stage the rank of the matrix equals the number of its rows.*

(3.1) *This process is continued until the set of all derivatives of the z's which are not rejected derivatives or derivatives of rejected derivatives are precisely the accepted derivatives.*

We shall prove in §5 that this must happen at some stage; and that, when it does, the α's involved in any minor of maximal rank at (x^0, α^0) are a complete set of parameters. The final matrix is denoted by (M).

The classical method of ending the algorithm is given in Theorem 5.9.

It is obvious that, if (3.1) is satisfied, any derivative of a rejected derivative cannot be an accepted derivative, and hence must itself be rejected if tested. (This simplification was surmised by S. S. Cairns.) For the present we cannot introduce this simplification.

Remark. It follows that each row of (M) except those determined by $\delta_1 f_1 \equiv f_1$, $\delta_1 f_2 \equiv f_2$, $\cdots$, $\delta_1 f_m \equiv f_m$ consists of derivatives of the elements of at least one earlier row with respect to some x_i.

4. **The system** (S). If N is the order of the highest derivative tested (with which the algorithm will eventually end), we assume that each f_i is of class C^{N+3} in (x, α) neighboring (x^0, α^0).

Suppose

$$Dz_s = \frac{\partial^{i_1+\cdots+i_n} z_s}{\partial x_1^{i_1} \cdots \partial x_n^{i_n}}$$

is a rejected derivative. We note that $Df_s(x, \alpha)$ is of class $C^{N+3-(i_1+\cdots+i_n)}$.

Let $D_1 z_{j_1}, D_2 z_{j_2}, \cdots, D_h z_{j_h}$ $(1 \leq j_k \leq r)$ be the derivatives accepted up to the time Dz_s is considered. (Note that the D's and j's are in general not all distinct.) Then the matrix thus far constructed is

$$\left\| \begin{matrix} \frac{\partial}{\partial \alpha_1}(D_1 f_{j_1}) & \cdots & \frac{\partial}{\partial \alpha_m}(D_1 f_{j_1}) \\ \cdots & \cdots & \cdots \\ \frac{\partial}{\partial \alpha_1}(D_h f_{j_h}) & \cdots & \frac{\partial}{\partial \alpha_m}(D_h f_{j_h}) \end{matrix} \right\|,$$

of rank h at (x^0, α^0). The trial row corresponding to Dz_s is

$$\frac{\partial}{\partial \alpha_1}(Df_s) \quad \frac{\partial}{\partial \alpha_2}(Df_s) \quad \cdots \quad \frac{\partial}{\partial \alpha_m}(Df_s).$$

Since Dz_s is rejected, this row does not increase the rank of the matrix at or near (x^0, α^0). Hence, by the general theorem of functional dependence, where $x_1, \cdots, x_n$ are considered as parameters and $\alpha_1, \cdots, \alpha_m$ as the independent variables, we infer that

$$(4.1) \qquad Df_s(x, \alpha) \underset{(x,\alpha)}{\equiv} H[x_1, \cdots, x_n, D_1 f_{i_1}(x, \alpha), \cdots, D_h f_{i_h}(x, \alpha)],$$

where $Df_s(x, \alpha)$ and H as a function of its $n + h$ arguments are known to be of class $C^{N+3-(i_1+\cdots+i_n)}$. In other words, $z_1 = f_1(x, \alpha), \cdots, z_r = f_r(x, \alpha)$ is a solution of the equation

$$(4.2) \qquad Dz_s = \frac{\partial^{i_1+\cdots+i_n} z_s}{\partial x_1^{i_1} \cdots \partial x_n^{i_n}} = H[x_1, \cdots, x_n, D_1 z_{i_1}, \cdots, D_h z_{i_h}].$$

Thus, whenever a derivative is rejected, there is a corresponding partial differential equation (ordinary if $n = 1$) similar to (4.2), satisfied by $z_1 = f_1(x, \alpha), \cdots, z_r = f_r(x, \alpha)$ in the neighborhood of (x^0, α^0). When the algorithm is complete, we have a system (S) of differential equations, whose left members are precisely the rejected derivatives, and whose right members are functions of the x's and accepted derivatives.

5. **Proof that the algorithm is finite.** For the system of differential equations built up at any stage, any derivative of a left member (including a 0-th derivative or left member itself) is called a *principal derivative*. All derivatives of $z_1, \cdots, z_r$ not principal derivatives of the system are called *parametric derivatives* of the system.

LEMMA 5.1. *Every accepted derivative is parametric for the system of differential equations built up at any stage of the algorithm.*

To prove this, let $\delta_e z_s$ be any accepted derivative, and (S_1) the system of differential equations of type (4.2) built up at any stage. Now if the highest left member in (S_1) is lower than $\delta_e z_s$, and (S_2) is the system of equations of type (4.2) having as left members all rejected derivatives lower than $\delta_e z_s$, then (S_2) includes all equations of (S_1). It then follows from the definition of "parametric" that $\delta_e z_s$ must be parametric for (S_1) if it is parametric for (S_2). Hence we may assume that the left members of (S_1) include all rejected derivatives lower than $\delta_e z_s$. All such are then principal derivatives of (S_1).

Suppose that $\delta_e z_s$ were principal for (S_1). Now $\delta_e z_s$ cannot be a left member of (S_1), for the left members of (S_1) are rejected derivatives, whereas $\delta_e z_s$ is accepted, by hypothesis. Hence it must be a derivative of a left member, that is, there must be an equation

$$(5.2) \qquad \delta_e z_s = H[x_1, \cdots, x_n, D_1 z_{i_1}, \cdots, D_t z_{i_t}]$$

of (S_1) such that

$$\delta_s z_s = \delta_c(\delta_e z_e),$$

where δ_c, like δ_s and δ_e, indicates certain differentiations to the x's. Now if $\delta_e z_e$ is a derivative of order p, we know that $\delta_e f_e$ exists, of class C^{N+3-p}. Hence $\delta_c(\delta_e f_e)$ exists and is of class C^{N+3-p}. The known class (in terms of its arguments) of the function on the right side of (5.2) is the same as the known class of $\delta_e f_e$; cf. (4.1). Since the differentiations δ_c can be performed on $\delta_e f_e$, they can therefore be performed on the right side of (5.2). Since (5.2) is satisfied by $z_1 = f_1(x, \alpha), \cdots, z_r = f_r(x, \alpha)$, the equation (E) obtained by performing differentiations δ_c on both sides of (5.2) is likewise satisfied by $z_1 = f_1(x, \alpha), \cdots, z_r = f_r(x, \alpha)$. Equation (E) has $\delta_s z_s$ as left member, and the function on the right side of (E) is of known class equal to the known class of $\delta_e f_e$, namely C^{N+3-p}.

Now if H actually has arguments other than $x_1, \cdots, x_n$, one or more new arguments will appear on the right in (E), namely some derivatives (G) of $D_1 z_{i_1}, \cdots, D_t z_{i_t}$. Since the derivatives on the right in (5.2) are lower than $\delta_e z_e$, the derivatives (G) are lower (not necessarily of lower order) than $\delta_s z_s$. The highest principal derivative in (G) can now be replaced by a function of $x_1, \cdots, x_n$ and derivatives lower than itself, by use of some equation of (S_1), differentiated if necessary, just as $\delta_s z_s$ was equated to the right side of (E) by use of (5.2); and the resulting equation is still satisfied by $z_1 = f_1(x, \alpha), \cdots, z_r = f_r(x, \alpha)$. In this way, step by step, all principal derivatives of (S_1) can be eliminated, since there can be only a finite number of derivatives in any descending sequence. Hence finally we have an equation

$$\delta_s z_s = \psi(x_1, \cdots, x_n, \delta_\sigma z_\beta, \cdots, \delta_\tau z_\gamma),$$

where $\delta_\sigma z_\beta, \cdots, \delta_\tau z_\gamma$ are parametric derivatives of (S_1) lower than $\delta_s z_s$, and

$$\delta_s f_s(x, \alpha) \underset{(x,\alpha)}{\equiv} \psi[x_1, \cdots, x_n, \delta_\sigma f_\beta(x, \alpha), \cdots, \delta_\tau f_\gamma(x, \alpha)]. \tag{5.3}$$

If a derivative $\delta_k z_l$ is replaced, under the procedure described above, in any equation

$$\delta_s z_s = K[x_1, \cdots, x_n, \cdots, \delta_k z_l, \cdots],$$

it is replaced by a function of known class at least as high as the known class of $\delta_s f_s$, since $\delta_k z_l$ is lower than $\delta_s z_s$. Hence the substitutions do not reduce the known class, and ψ must be of the same known class as that of the right member of (E), namely, C^{N+3-p}.

Since any derivative lower than $\delta_s z_s$ has been either accepted or rejected before $\delta_s z_s$ is considered, and the rejected derivatives lower than $\delta_s z_s$ are principal derivatives of (S_1), the (parametric) derivatives $\delta_\sigma z_\beta, \cdots, \delta_\tau z_\gamma$ must be accepted derivatives. Hence $\delta_\sigma f_\beta, \cdots, \delta_\tau f_\gamma$ are among the functions whose derivatives with respect to $\alpha_1, \cdots, \alpha_m$ appear in the first $w - 1$ rows of the matrix, if we

suppose $\delta_s z_s$ to be the w-th accepted derivative. For fixed (x) near (x^0) we then infer, from (5.3), by the general theorem of functional dependence, that the trial row

$$\frac{\partial}{\partial \alpha_1}(\delta_s f_s) \qquad \frac{\partial}{\partial \alpha_2}(\delta_s f_s) \qquad \cdots \qquad \frac{\partial}{\partial \alpha_m}(\delta_s f_s)$$

cannot increase the rank of the matrix for (α) near (α^0). Hence $\delta_s z_s$ must be a rejected derivative, contrary to hypothesis. We infer that Lemma 5.1 is true.

COROLLARY 5.4. *The derivatives on the right in any equation* (4.2) *are parametric for the system of such equations built up at any stage of the algorithm.*

For, every derivative on the right in such an equation is an accepted derivative.

LEMMA 5.5. *If, under the algorithm, all partial derivatives of a given order A are rejected, and the algorithm is incomplete before the last of them is considered, then its rejection completes the algorithm.*

We begin by rewriting condition (3.1) in the following obviously equivalent form:

(5.6) *The algorithm is continued till the parametric derivatives of the system of differential equations are precisely the accepted derivatives.*

Now, suppose the hypotheses of the lemma are satisfied. Denote by (S_1) the system of differential equations which has been constructed when the last derivative of order A has been rejected. Since all the derivatives of order A become left members in (S_1), we see that all derivatives of higher orders must be principal for (S_1). Hence all parametric derivatives must be of lower orders. The rejected derivatives of lower orders are left members in (S_1), hence not parametric. Therefore the parametric derivatives must be accepted derivatives. But, by Lemma 5.1, all accepted derivatives are parametric. Hence the accepted derivatives are precisely the parametric derivatives, so that (5.6) is satisfied, and the proof of Lemma 5.5 is complete. Note that (S_1) is therefore the (S) of §4.

THEOREM 5.7. *The algorithm is finite.*

Since the rank of the matrix cannot exceed the number m of its columns, a stage must be reached after which all derivatives tested are rejected. Denote by T the order of the highest accepted derivative. Then, if the algorithm does not end before testing the last derivative of order $A = T + 1$, it must, by Lemma 5.5, end when the last of them is tested (and rejected). Hence the theorem is true.

LEMMA 5.8. *If the f's are of class C^E, $E \geq N + 3$, where N is the order of the final derivative tested under* (3.1), *and the algorithm is continued after condition* (3.1) [*or* (5.6)] *has been satisfied, but where no derivative of the f's with respect to x's of*

order higher than $E - 1$ has been tested, all derivatives tested after (3.1) *is satisfied will be rejected, so that* (M) *will not become larger.*

By hypothesis the highest derivative to be tested is of order $E - 1 \geq N + 2$. Now, while in the proof of Lemma 5.1, $p \leq N$, the entire proof of Lemma 5.1 goes through for the present case, where $p \leq N + 2$. (The functions in the proof of Lemma 5.1 of class C^{N+3-p} are now of class C^1 at least. It is not until §7 is reached that higher class is needed.) Hence if another derivative, say w, were accepted, it would be parametric for the system of differential equations at any stage, in particular for (S). Since this would contradict the fact that the parametric derivatives for (S) are precisely the derivatives accepted up to the time (3.1) was satisfied and hence do not include w, we infer that Lemma 5.8 is true.

The following theorem establishes the equivalence of our test to the classical test, under which all rows remain in the matrix and the algorithm is continued until all derivatives of some order fail to increase the rank. Since in the classical case the rank of the matrix equals the minimum number of essential parameters by means of which the given function can be expressed with identities in (x, α), we infer that that number equals the number of parameters in a complete set of effective parameters, at any non-singular point. (Cf. Theorem 8.1.)

THEOREM 5.9. *If instead of using conditions* (3.1) *or* (5.6) *we continue the algorithm until for the first time all derivatives of a given order A are rejected, then the matrix* (M) *will be complete. At the same time, in testing derivatives for* (M), *any derivative of a rejected derivative can be rejected without trial.*

The final sentence of this theorem was surmised by S. S. Cairns.

Proof. From Lemma 5.5 we see that the highest accepted derivative must be of order at least $N - 1$, where N is the order of the highest derivative tested under (3.1) or (5.6).

Applying Lemma 5.8 with $E = N + 3$, we see that the situation called for in the first part of Theorem 5.9 must occur, with $A = N$ or $N + 1$ according as the highest accepted derivative is of order $N - 1$ or N. Lemma 5.5 shows that, the first time it does occur, (3.1) will be satisfied by the time the last derivative of order A has been rejected, if not before. Hence (M) must be complete.

As for the final conclusion of the theorem, we see by (3.1) that a derivative of a rejected derivative cannot be an accepted derivative and hence must itself be rejected if tested. This completes the proof.

Remark. From Theorem 5.9 we infer that if the algorithm is carried through in this way, the ordering of the derivatives is immaterial, except that a derivative of lower order should precede one of higher order.

6. **The system** (T). In this section we obtain a system (T) of differential equations, determined by the system (S) and the matrix (M), which will enable

us to prove the existence of our complete set of effective parameters. As (T) is built up, we shall see that at each stage it is satisfied by $z_1 = f_1(x, \alpha), \cdots, z_r = f_r(x, \alpha)$, when certain quantities p_i are replaced by derivatives of the f's which will correspond to them.

Suppose the number of parametric derivatives (number of rows of (M)) is k, and let these derivatives be $D_1z_{s_1}, D_2z_{s_2}, \cdots, D_kz_{s_k}$, in their order as defined in §3. (Here operators D_w and D_y may be the same even if $w \neq y$, and s_a may equal s_b even if $a \neq b$.) We introduce the symbols $p_1, p_2, \cdots, p_k$, and shall say that p_i *corresponds to* $D_iz_{s_i}$ $(i = 1, 2, \cdots, k)$. The derivative δp_i will be said to *correspond to* the derivative of z_{s_i} given by $\delta(D_iz_{s_i})$. The dependent variables of our system (T) will be $p_1, p_2, \cdots, p_k$.

For any $i \leq k$ and $j \leq n$, $\partial p_i/\partial x_j$ corresponds to either a principal or a parametric derivative of (S). If it corresponds to a parametric derivative, say D_lz_s, then the equation

$$\frac{\partial p_i}{\partial x_j} = p_l \tag{6.1}$$

is put into (T). If it corresponds to a principal derivative v which is a left member of (S), then the equation of (S) with v as left member is put into (T), but with v replaced by $\partial p_i/\partial x_j$, and all (parametric) derivatives on the right replaced by the corresponding p's. To be sure, the equation of (S) having v as left member may be used more than once in this way.

The remaining possibility is that $\partial p_i/\partial x_j$ corresponds to a principal derivative of (S) which is not a left member, but a derivative δ of a left member v of (S). (To make v unique, let us take it as low as possible.) In this case we take the derivative δ of both sides of the equation containing v, which may give rise to certain principal derivatives on the right side. But these principal derivatives are lower than δv, since differentiation preserves relative order. Hence, by use of other equations of (S), differentiated or not, we can replace these principal derivatives by functions of $x_1, \cdots, x_n, D_1z_{s_1}, \cdots, D_kz_{s_k}$, and possibly still other—but lower—principal derivatives. By a finite number of such steps all principal derivatives on the right side can be eliminated. The resulting equation, with the left member replaced by $\partial p_i/\partial x_j$, and all derivatives on the right replaced by the corresponding p's, is put into (T).

Thus finally we obtain a system (T) of $k \cdot n$ differential equations:

$$\frac{\partial p_i}{\partial x_j} = H_{ij}(x_1, \cdots, x_n, p_1, \cdots, p_k) \qquad (i = 1, \cdots, k; j = 1, \cdots, n). \tag{6.2}$$

At each stage, as at the start (cf. (4.1)), the right side of an equation of (S), or of a differentiated equation of (S), has class $N + 3$ minus the order of the derivative on the left. Furthermore, the substitutions mentioned above do not disturb this situation, since any replaced derivative is never of higher order than

the left member of any equation containing it. Hence a similar statement is true for (T). Since the highest derivative corresponding to a left member of (6.2) is of order at most $N + 1$, we see that all H_{ij} in (6.2) are of class $C^{N+3-(N+1)}$, or C^2 (and possibly higher).

If $F_i(x, \alpha) \equiv D_i f_{s_i}(x, \alpha)$ $(i = 1, \cdots, k)$, every F_i is of class C^3.

LEMMA 6.3. *For any* (α) *near* (α^0), *the equations*

$$p_i = F_i(x, \alpha) \qquad (i = 1, \cdots, k) \tag{6.4}$$

yield a solution of (6.2) *for* (x) *near* (x^0).

For, since $z_1 = f_1(x, \alpha), \cdots, z_r = f_r(x, \alpha)$ is a solution of (S), and (T) is derived from (S), $z_1 = f_1, \cdots, z_r = f_r$ satisfies the system obtained from (T) by substituting for the p's the corresponding derivatives of the z's. Hence (6.4) satisfy (T), and the lemma is valid.

7. **Proof that (T) are integrable, and some consequences.** Using the notation of §6, let

$$p_i^0 = F_i(x^0, \alpha^0) \qquad (i = 1, \cdots, k). \tag{7.1}$$

LEMMA 7.2. *The integrability conditions for* (6.2) *are identically satisfied, for* (x, p) *in a neighborhood of* (x^0, p^0).

While the writer has not seen a treatment of a system like (6.2) with the H_{ij} non-analytic, the theory is similar to the case $k = 1$. Cf. [2]. That the H_{ij} be of class C^2 is sufficient for the existence and uniqueness of solutions, provided the integrability conditions are satisfied:

$$(\mathfrak{s})\quad \frac{\partial H_{ij}}{\partial x_l} + \sum_{s=1}^{k} \frac{\partial H_{ij}}{\partial p_s}\cdot H_{sl} \equiv \frac{\partial H_{il}}{\partial x_j} + \sum_{s=1}^{k} \frac{\partial H_{il}}{\partial p_s}\cdot H_{sj}$$

$$(i = 1, \cdots, k; j, l = 1, \cdots, n).$$

Proof. Since the F's are exactly the derivatives of the f's to x's figuring in (M), which is of rank k at (x^0, α^0), some Jacobian of derivatives of the F's with respect to k of the α's is not zero at (x^0, α^0), say

$$\frac{D(F_1, \cdots, F_k)}{D(\alpha_1, \cdots, \alpha_k)} \neq 0 \qquad \text{at } (x^0, \alpha^0). \tag{7.3}$$

Hence, in some neighborhood of (x^0, α^0, p^0), equations (6.4) are equivalent to a system of the form

$$\alpha_i = \phi_i(x_1, \cdots, x_n, p_1, \cdots, p_k, \alpha_{k+1}, \cdots, \alpha_m) \quad (i = 1, \cdots, k), \tag{7.4}$$

with the ϕ_i's of class C^3 since the right members of (6.4) are of class C^3.

If (x^1, p^1) is given near (x^0, p^0), then $\alpha_1^1, \cdots, \alpha_k^1$ are uniquely determined by the equations

$$\alpha_i^1 = \phi_i(x_1^1, \cdots, x_n^1, p_1^1, \cdots, p_k^1, \alpha_{k+1}^0, \cdots, \alpha_m^0) \qquad (i = 1, \cdots, k).$$

From the equivalence of (6.4) and (7.4) we see that

$$p_i^1 = F_i(x_1^1, \cdots, x_n^1, \alpha_1^1, \cdots, \alpha_k^1, \alpha_{k+1}^0, \cdots, \alpha_m^0) \qquad (i = 1, \cdots, k). \tag{7.5}$$

Since (6.2) are satisfied by (6.4), they are satisfied by

$$p_i = F_i(x_1, \cdots, x_n, \alpha_1^1, \cdots, \alpha_k^1, \alpha_{k+1}^0, \cdots, \alpha_m^0) \qquad (i = 1, \cdots, k). \tag{7.6}$$

Hence (7.6) must satisfy the integrability conditions $(\mathcal{J})$ of (6.2), equations in in (x, p). In particular, (7.6) satisfy $(\mathcal{J})$ at the point on (7.6) where $(x) = (x^1)$, and comparing with (7.5) we see that $(\mathcal{J})$ are satisfied at (x^1, p^1). Since the latter was any point neighboring (x^0, p^0), we infer that $(\mathcal{J})$ are identically satisfied, and the lemma is proved.

LEMMA 7.7. *Suppose (M) is of rank k and* (7.3) *holds. The solution of* (6.2) *determined by*

$$p_i = p_i^1 \quad \text{at } (x) = (x^0) \qquad (i = 1, \cdots, k) \tag{7.8}$$

for any $(p_1^1, \cdots, p_k^1)$ in a neighborhood $\mathcal{P}$ of $(p_1^0, \cdots, p_k^0)$ is

$$p_i = F_i(x_1, \cdots, x_n, \beta_1, \cdots, \beta_k, \alpha_{k+1}^0, \cdots, \alpha_m^0) \qquad (i = 1, \cdots, k), \tag{7.9}$$

with

$$\beta_i = \phi_i(x_1^0, \cdots, x_n^0, p_1^1, \cdots, p_k^1, \alpha_{k+1}^0, \cdots, \alpha_m^0) \qquad (i = 1, \cdots, k), \tag{7.10}$$

where the ϕ's are the functions appearing in (7.4).

Proof. Since (6.4) are equivalent to (7.4), we infer from (7.10) that

$$p_i^1 = F_i(x_1^0, \cdots, x_n^0, \beta_1, \cdots, \beta_k, \alpha_{k+1}^0, \cdots, \alpha_m^0) \qquad (i = 1, \cdots, k). \tag{7.11}$$

By Lemma 6.3, (7.9) yields a solution of (6.2). From (7.11) we see that this solution satisfies (7.8). Since (7.8) determines just one solution, (7.9) must be it, and the lemma is proved.

LEMMA 7.12. *Suppose (M) is of rank k and* (7.3) *holds. Given any neighborhood K of $(\alpha_1^0, \cdots, \alpha_k^0)$ in k-space, there corresponds a neighborhood $\mathfrak{M}$ of (α^0) in m-space such that the totality of k-tuples of functions of $x_1, \cdots, x_n$ for (x) near (x^0) obtained from $F_1(x, \alpha), \cdots, F_k(x, \alpha)$ with all (α) in $\mathfrak{M}$ is included in the totality obtained by taking $\alpha_{k+1} = \alpha_{k+1}^0, \cdots, \alpha_m = \alpha_m^0$; and $(\alpha_1, \cdots, \alpha_k)$ in K.*

Since (β) and (p^1) are in one-to-one continuous correspondence, in the neighborhood of $(\alpha_1^0, \cdots, \alpha_k^0)$ and (p^0), respectively, under (7.10), to the given neighborhood K of $(\alpha_1^0, \cdots, \alpha_k^0)$ corresponds a neighborhood $\mathcal{P}$ of (p^0), such that if (p) is in $\mathcal{P}$, then (β) will be in K. We choose $\mathcal{P}$ sufficiently small to be used in Lemma 7.7. Now let $\mathfrak{M}$ be a neighborhood of $(\alpha_1^0, \cdots, \alpha_m^0)$ in m-space such that if (α) is in $\mathfrak{M}$, then $[F_1(x^0, \alpha), \cdots, F_k(x^0, \alpha)]$ will be in $\mathcal{P}$.

We take $\mathfrak{M}$ sufficiently small to insure that if (α^1) is in $\mathfrak{M}$ we can apply Lemma 6.3, and infer that $p_1 = F_1(x, \alpha^1), \cdots, p_k = F_k(x, \alpha^1)$ is a solution of (6.2). For this solution, when $(x) = (x^0)$, $p_i = F_i(x^0, \alpha^1)$, which we denote by p_i^1 $(i = 1, \cdots, k)$. According to the last sentence of the preceding paragraph, (p^1) is then in $\mathcal{P}$. By Lemma 7.7, the functions in this solution are the same k-tuple of functions of $(x_1, \cdots, x_n)$ as the $F_1, \cdots, F_k$ given by (7.9) under conditions (7.10). Furthermore, the (β) of (7.10) is in K, by the first sentence of the preceding paragraph. Since in (7.9) the last $m - k$ parameters are $\alpha_{k+1}^0, \cdots, \alpha_m^0$, the lemma is proved.

LEMMA 7.13. *If (M) is of rank k and (7.3) holds, the set $\alpha_1, \cdots, \alpha_k$ forms a sufficient set of parameters for $\{f(x, \alpha)\}$ at (x^0, α^0).*

Proof. Among the functions $f_1, \cdots, f_r$ suppose $f_{a_1}, \cdots, f_{a_h}$ are accepted (zeroth) derivatives, and the rest, say $f_{b_1}, \cdots, f_{b_{r-h}}$, are rejected. Then $f_{a_1}, \cdots, f_{a_h}$ are included among $F_1, \cdots, F_k$, and, as is shown in §4, $f_{b_1}, \cdots, f_{b_{r-h}}$ are functions of $x_1, \cdots, x_n, F_1, \cdots, F_k$:

$$f_{b_i}(x, \alpha) \underset{(x,\alpha)}{\equiv} \zeta_i[x_1, \cdots, x_n, F_1(x, \alpha), \cdots, F_k(x, \alpha)] \qquad (i = 1, \cdots, r - h). \tag{7.14}$$

Let K and $\mathfrak{M}$ be as in Lemma 7.12, and let any (α) in $\mathfrak{M}$ be given. By Lemma 7.12 the k-tuple $F_1(x, \alpha), \cdots, F_k(x, \alpha)$ is the same k-tuple of functions of (x) as the k-tuple

$$F_1(x; \beta_1, \cdots, \beta_k, \alpha_{k+1}^0, \cdots, \alpha_m^0), \cdots, F_k(x; \beta_1, \cdots, \beta_k, \alpha_{k+1}^0, \cdots, \alpha_m^0),$$

for some $(\beta_1, \cdots, \beta_k)$ in K. Hence

$$f_{a_j}(x; \beta_1, \cdots, \beta_k, \alpha_{k+1}^0, \cdots, \alpha_m^0) \underset{(x)}{\equiv} f_{a_j}(x; \alpha_1, \cdots, \alpha_m) \qquad (j = 1, \cdots, h), \tag{7.15}$$

and, using (7.14),

$$f_{b_i}(x; \beta_1, \cdots, \beta_k, \alpha_{k+1}^0, \cdots, \alpha_m^0)$$

$$\underset{(x,\beta)}{\equiv} \zeta_i[x_1, \cdots, x_n, F_1(x; \beta_1, \cdots, \beta_k, \alpha_{k+1}^0, \cdots, \alpha_m^0),$$

$$
\begin{aligned}
(7.16) \qquad & \cdots , F_k(x; \beta_1 , \cdots , \beta_k , \alpha^0_{k+1} , \cdots , \alpha^0_m)] \\
& \underset{(x)}{\equiv} \zeta_i[x_1 , \cdots , x_n , F_1(x; \alpha_1 , \cdots , \alpha_m), \cdots , F_k(x; \alpha_1 , \cdots , \alpha_m)] \\
& \underset{(x)}{\equiv} f_{b_i}[x; \alpha_1 , \cdots , \alpha_m] \qquad (i = 1, \cdots , r - h).
\end{aligned}
$$

From (7.15) and (7.16) we infer that

$$
f_l(x; \beta_1 , \cdots , \beta_k , \alpha^0_{k+1} , \cdots , \alpha^0_m) \underset{(x)}{\equiv} f_l[x; \alpha_1 , \cdots , \alpha_m] \qquad (l = 1, \cdots , r).
$$

Hence the first k parameters form a sufficient set, and the lemma is proved.

8. **The principal results.** The principal theorem follows.

THEOREM 8.1. *Given* $f_1(x_1 , \cdots , x_n , \alpha_1 , \cdots , \alpha_m), \cdots , f_r(x_1 , \cdots , x_n , \alpha_1 , \cdots , \alpha_m)$ *of class specified below in the neighborhood of* (x^0, α^0), *we suppose that* (x^0, α^0) *is not a singular point, and that the matrix* (M) *is constructed, the conclusion of the algorithm being determined either as in* (3.1) *or as in Theorem* 5.9. *If* N *is the order of the last derivative tested, we assume that the* f's *are of class* C^{N+3} *in the variables* (x, α). *Let* k *be the number of rows in* (M), *and suppose* $\alpha_{i_1} , \cdots , \alpha_{i_k}$ *are parameters figuring in a* k*-rowed determinant of* (M) *which is not zero at* (x^0, α^0). *Then* $\alpha_{i_1} , \cdots , \alpha_{i_k}$ *form a complete set of parameters for* $\{f(x, \alpha)\}$ *at* (x^0, α^0).

A third method for ending the algorithm is to continue till a sequence (W) of derivatives has been rejected such that all derivatives of the z's higher than the derivatives in (W) are derivatives of derivatives in (W).

Without loss of generality we may assume that $(\alpha_{i_1} , \cdots , \alpha_{i_k})$ is $(\alpha_1 , \cdots , \alpha_k)$, so that (7.3) holds. Hence the equations

$$
u_i = F_i(x^0_1 , \cdots , x^0_n , \alpha_1 , \cdots , \alpha_k , \alpha^0_{k+1} , \cdots , \alpha^0_m) \qquad (i = 1, \cdots , k)
$$

place $(\alpha_1 , \cdots , \alpha_k)$ in one-to-one correspondence with $(u_1 , \cdots , u_k)$, locally. Hence two non-identical k-tuples $(\alpha^1_1 , \cdots , \alpha^1_k)$ and $(\alpha^2_1 , \cdots , \alpha^2_k)$ determine two non-identical k-tuples $(u^1_1 , \cdots , u^1_k)$ and $(u^2_1 , \cdots , u^2_k)$, and two non-identical k-tuples of values for $\{F(x^0; \alpha^1_1 , \cdots , \alpha^1_k , \alpha^0_{k+1} , \cdots , \alpha^0_m)\}$ and $\{F(x^0; \alpha^2_1 , \cdots , \alpha^2_k , \alpha^0_{k+1} , \cdots , \alpha^0_m)\}$. Since the F's are derivatives of some of the f's with respect to the x's, it follows that $\{f(x; \alpha^1_1 , \cdots , \alpha^1_k , \alpha^0_{k+1} , \cdots , \alpha^0_m)\}$ cannot then be the same k-tuple of functions of (x) as $\{f(x; \alpha^2_1 , \cdots , \alpha^2_k , \alpha^0_{k+1} , \cdots , \alpha^0_m)\}$. Hence $\alpha_1 , \cdots , \alpha_k$ are effective parameters for $\{f\}$ at (x^0, α^0). By Lemma 7.13 they are also sufficient. Hence, they form a complete set of parameters, and the theorem is proved.

REMARK. *We note that the proof that the parameters are effective does not require the* f's *to be of class as high as is required in Theorem* 8.1, *but simply that the derivatives occurring in* (7.3) *be continuous.*

The writer has proved the following theorem, but will omit the proof, partly because of the similarity to Theorem 10.5.

THEOREM 8.2. *Under the hypotheses of Theorem* 8.1, *no smaller number than* k *of the given parameters can be sufficient, and no larger number of them can be effective.*

The hypotheses of this theorem do not rule out singular points in the case of the smaller or larger number of parameters.

9. **The singular points.** The following theorem in the large shows that the restriction to non-singular points is not as serious as might be imagined.

THEOREM 9.1. *Let* $\mathfrak{R}$ *be an open set in* (x, α)*-space in which* $f_1(x_1, \cdots, x_n, \alpha_1, \cdots, \alpha_m), \cdots, f_r(x_1, \cdots, x_n, \alpha_1, \cdots, \alpha_m)$ *are sufficiently differentiable so that each point of* $\mathfrak{R}$ *can be established as singular or as non-singular with the conditions of Theorem* 8.1 *satisfied. The locus of singular points in* (x, α)*-space is then a nowhere dense set, closed relative to* $\mathfrak{R}$.

Proof. Since the rank of (M) cannot exceed m, a finite number of steps will suffice to determine whether a given point is singular or non-singular. From the definition it is then seen that, if (x^0, α^0) is a non-singular point, all points in a neighborhood of (x^0, α^0) are non-singular. Hence the locus of non-singular points is open, and its complement, the locus E of singular points, is closed relative to $\mathfrak{R}$.

We infer that, if E were dense in any region, the region would consist of points of E. But consideration of the algorithm for the construction of (M) shows that in any neighborhood of a singular point a non-singular point can be found; hence there can be no region of singular points. This contradiction shows that E is nowhere dense, and the theorem is proved.

REMARK 9.2. *If the order of the* x*'s is changed (or what is equivalent, the ordering of the derivatives is changed to what it would be under a change of order in the* x*'s), a singular point may become non-singular, or vice versa.*

This is shown by the example $f(x, y, \alpha, \beta) \equiv \alpha x + \beta y$, with $(x^0, y^0) = (1, 0)$ and α^0, β^0 arbitrary, $r = 1$. We omit details.

REMARK 9.3. *If the order of the* f*'s is changed, a singular point may become non-singular, or vice versa.*

To show this, consider the pair of functions $\alpha + x$ and αx. If $f_1 = \alpha + x$ and $f_2 = \alpha x$, then the matrix (M) is $\| 1 \|$, and there are no singular points. If $f_1 = \alpha x$ and $f_2 = \alpha + x$, the first trial row of (M) consists of the element x, which is zero if $x = 0$. Hence in this case any point at which $x = 0$ is a singular point.

THEOREM 9.4. *The numbers of parameters in complete sets for different non-singular points neighboring a given singular point may be unequal. However, if the* f*'s are analytic in* (x, α) *they must be equal.*

First, to show that the numbers may be unequal, consider

$$f(x, \alpha) \equiv \begin{cases} 0, & x \leq 0, \\ \\ \alpha x^3, & x > 0 \end{cases} \qquad (r = 1).$$

The singular points are those at which $x = 0$. At any point where $x < 0$, the number of parameters in a complete set is zero. At any point where $x > 0$, the number is one.

If every $f_i(x, \alpha)$ is analytic, the singular points neighboring (x^0, α^0) in the space of the complex variables $x_1, \cdots, x_n, \alpha_1, \cdots, \alpha_m$ are easily seen to be confined to a finite number of loci obtained by equating to zero analytic functions which are not identically zero. Hence the locus of singular points cannot separate any two non-singular points in any neighborhood of (x^0, α^0) in the space of the complex variables. As the number of parameters in a complete set is constant in a neighborhood of any non-singular point, it is therefore the same for all non-singular points neighboring (x^0, α^0).

We now give an example of an analytic function to show that a singularity may be a result not merely of the definition of singular point forced on us by the algorithm used but of our definitions of effective and sufficient. Let

$$(9.5) \qquad f(x, \alpha, \beta) = \alpha\beta + x \qquad (r = 1).$$

The first trial row of the matrix (M) is then

$$\beta \qquad\qquad \alpha.$$

Hence the point $(x^0, 0, 0)$ is a singular point for any x^0. At $(x^0, 0, 0)$ neither α nor β is sufficient, and the set (α, β) is not effective, as is easily verified from (9.5). Hence there is no complete set of parameters for that point.

In the case of this example, the classical treatment is clearly superior, since under it we would consider the family $F(A, x) \equiv A + x$. It is in the (general) case that F is not obvious that the treatment of this paper will be advantageous.

10. **The number of parameters needed.** In this section we show that no smaller number of parameters than that in a complete set will suffice.

LEMMA 10.1. *Suppose we are given* $\{f(x_1, \cdots, x_n, \alpha_1, \cdots, \alpha_m)\}$ *and* $\{g(x_1, \cdots, x_n, \beta_1, \cdots, \beta_l)\}$, *r-tuples of functions satisfying the hypotheses of Theorem* 8.1 *neighboring* (x^0, α^0) *and* (x^0, β^0) *respectively; suppose that the g's are of class one greater than the N for* $\{f\}$. *If, for* (x) *in some neighborhood* $\mathfrak{N}$ *of* (x^0) *in* (x)*-space, the r-tuples* $\{g(x, \beta)\}$ *of functions of* (x) *for* (β) *in any given neighborhood* $\mathfrak{G}$ *of* (β^0) *include all r-tuples* $\{f(x, \alpha)\}$ *of functions of* (x) *for* (α) *in a corresponding neighborhood* $\mathfrak{F}$ *of* (α^0), *the number of parameters in a complete set for* $\{g(x, \beta)\}$ *at* (x^0, β^0) *equals or exceeds the number of parameters in a complete set for* $\{f(x, \alpha)\}$ *at* (x^0, α^0).

It is to be noted that we do not assume that the parameters of either set can be expressed as functions of the parameters of the other set.

It will be sufficient to show that every derivative that is accepted for $\{f\}$ is likewise accepted for $\{g\}$.

Suppose this is true up to a certain point; let δz_p be the next accepted derivative for $\{f\}$, if any. Now if δz_p were not accepted for $\{g\}$ it must be either rejected or not tested for $\{g\}$. According to the hypothesis on the class of the g's, we *could* test δz_p for $\{g\}$ even if the (M) for $\{g\}$ were complete, and, by Lemma 5.8, δz_p would then be rejected for $\{g\}$. We may therefore assume that, if δz_p were not accepted for $\{g\}$, it would be rejected for $\{g\}$.

Let $D_1 z_{s_1}, D_2 z_{s_2}, \cdots, D_h z_{s_h}$ be the derivatives accepted for $\{g\}$ up to the time of testing δz_p for $\{g\}$. Let $F_i(x, \alpha) \equiv D_i f_{s_i}(x, \alpha)$, and $G_i(x, \beta) \equiv D_i g_{s_i}(x, \beta)$ $(i = 1, \cdots, h)$.

Since δz_p would be rejected for $\{g(x, \beta)\}$, we would have

$$\delta g_p(x, \beta) \equiv K[x_1, \cdots, x_n, G_1(x, \beta), \cdots, G_h(x, \beta)], \tag{10.2}$$

for (x, β) neighboring (x^0, β^0), with K of class C^1 (at least) in its $n + h$ arguments.

Let $\mathcal{G}$ be a neighborhood of (β^0) such that if (β) is in $\mathcal{G}$ and (x) is in a neighborhood $\mathcal{N}_1$ of (x^0), then (10.2) holds. We take $\mathcal{N}_1 \subset \mathcal{N}$. Let $\mathcal{F}$ correspond to $\mathcal{G}$ as in the hypotheses of the theorem. For any (α^1) in $\mathcal{F}$, let (β^1) be a point in $\mathcal{G}$ such that

$$f_j(x, \alpha^1) \underset{(x)}{\equiv} g_j(x, \beta^1) \qquad (j = 1, \cdots, r),$$

for (x) in $\mathcal{N}$, hence also in $\mathcal{N}_1$. Then

$$F_i(x, \alpha^1) \underset{(x)}{\equiv} D_i f_{s_i}(x, \alpha^1) \underset{(x)}{\equiv} D_i g_{s_i}(x, \beta^1) \underset{(x)}{\equiv} G_i(x, \beta^1) \qquad (i = 1, \cdots, h).$$

Also,

$$\delta f_p(x, \alpha^1) \underset{(x)}{\equiv} \delta g_p(x, \beta^1).$$

Substituting in (10.2) with $(\beta) = (\alpha^1)$, we would have, for (x) in $(\mathcal{N}_1)$,

$$\delta f_p(x, \alpha^1) \equiv K[x_1, \cdots, x_n, F_1(x, \alpha^1), \cdots, F_h(x, \alpha^1)]. \tag{10.3}$$

Since (10.3) would hold for any (α^1) in $\mathcal{F}$ we would thus have

$$\delta f_p(x, \alpha) \underset{(x, \alpha)}{\equiv} K[x_1, \cdots, x_n, F_1(x, \alpha), \cdots, F_h(x, \alpha)]. \tag{10.4}$$

By the supposition above, $F_1, \cdots, F_h$ include all the functions figuring in the part of (M) thus far constructed for $\{f(x, \alpha)\}$, as well as perhaps some which are functions of $x_1, \cdots, x_n$ and those figuring in that part. Hence, for any fixed (x) in $\mathcal{N}_1$, (10.4) would show, by the general theorem of functional dependence, that, when we take the trial row consisting of the derivatives of $\delta f_p(x, \alpha)$ to $\alpha_1, \alpha_2,$

$\cdots, \alpha_m$, in the matrix (M) for $\{f(x, \alpha)\}$, the rank would not be increased. Hence δz_p would be rejected for $\{f(x, \alpha)\}$, contrary to hypothesis. We infer the truth of Lemma 10.1.

THEOREM 10.5. *Suppose the hypotheses of Theorem* 8.1 *are satisfied by* $\{f(x_1, \cdots, x_n, \alpha_1, \cdots, \alpha_m)\}$ *and* $\{g(x_1, \cdots, x_n, \beta_1, \cdots, \beta_t)\}$ *neighboring* (x^0, α^0) *and* (x^0, β^0), *respectively, and that the g's are of class one greater than the N for* $\{f\}$. *If, for* (x) *in a neighborhood of* (x^0), *the r-tuple* $\{g(x; \beta_1, \cdots, \beta_t)\}$ *of functions of* (x) *for* (β) *in any given neighborhood of* (β^0) *include all r-tuples* $\{f(x; \alpha_1, \cdots, \alpha_m)\}$ *of functions of* (x) *for* (α) *in a corresponding neighborhood of* (α^0), *then* $t \geq$ *the number k of parameters in a complete set for* $\{f\}$.

Proof. According to Lemma 10.1, the number of parameters in a complete set for $\{g\} \geq k$. Since it necessarily $\leq t$, we infer that $k \leq t$, as was to be proved.

BIBLIOGRAPHY

1. L. BIANCHI, *Lezione sulla teoria dei gruppi continui finiti di transformationi*, Pisa, 1918, pp. 34–37.
2. LUDWIG BIEBERBACH, *Differentialgleichungen*, third edition, p. 299.
3. L. P. EISENHART, *Continuous groups of transformations*, Princeton University Press, 1933, pp. 9–12.
4. J. F. RITT, *Algebraic differential equations*, American Mathematical Society, Colloquium Publications, New York, 1932, Chapter 9.

QUEENS COLLEGE.

ASYMPTOTIC DEVELOPMENTS OF CERTAIN INTEGRAL FUNCTIONS

By Cleota G. Fry and H. K. Hughes

I. Theorems of Ford and Newsom

1. **Introduction.** In several papers extending over the years 1904-08, E. W. Barnes (see especially [1]) determined the asymptotic behavior in the neighborhood of the point at infinity of a number of analytic functions defined by their Maclaurin developments. Moreover, several investigations of a similar nature have been made very recently by E. M. Wright [5], [6], [7]. The work of Barnes and Wright, in each instance, consists largely of a detailed study of the particular function considered. On the other hand, W. B. Ford [2; 4–15, 30–37] and C. V. Newsom [4] have recently established certain theorems which are general in character and which may be applied to a variety of different functions. In fact, they may be used to obtain the asymptotic developments of several of the specific functions considered by Barnes and Wright.

The present paper presents an application of the theorems of Ford and Newsom. A certain extension of Newsom's theorem is first stated, the proof being omitted. We then proceed to determine the asymptotic developments of the general integral function

$$F_\alpha(z) = \sum_{n=0}^{\infty} \frac{h(n)z^n}{\Gamma(\alpha n + t)} \qquad (\alpha > 0), \tag{1.1}$$

where t is any constant, real or complex, and where the function $h(n)$ depends only on n and satisfies certain further conditions. This work constitutes Part II of the paper. In Part III, we apply the theorem obtained in Part II to the special function

$$E_\alpha(z, \theta, \beta) = \sum_{n=0}^{\infty} \frac{z^n}{(n + \theta)^\beta\, \Gamma(\alpha n + 1)} \qquad (\alpha > 0), \tag{1.2}$$

where θ and β are any constants, real or complex, except that θ cannot equal zero or a negative integer. The asymptotic developments of both the functions given by (1.1) and (1.2) have been discussed by Ford for the special case in which $\alpha = 1$. His results are, in fact, a special case of those obtained in Parts II and III of this paper.

The most recent work on the function (1.1) appears to be that of Wright [5]. We shall refer to it later. Barnes has investigated the function (1.2) under the condition that θ is not an integer.

Received April 9, 1942.

2. **Statement of theorems.** The general results of Ford and Newsom cited in §1 may, for our present purpose, be put in the form of two fundamental theorems which we now state without proof. These theorems provide the starting point of our investigations.

THEOREM I. *Suppose that the coefficient $g(n)$ of the power series*

$$f(z) = \sum_{n=0}^{\infty} g(n)z^n \tag{2.1}$$

with infinite radius of convergence may be considered as a function $g(w)$ of the complex variable $w = x + iy$ and as such satisfies the following two conditions when considered throughout any arbitrary right half-plane $x > x_0$:

(a) *it is single-valued and analytic;*

(b) *for all values of $| y |$ sufficiently large, we may write*

$$| g(x + iy) | < Ke^{(\gamma\pi+\epsilon)|y|}, \tag{2.2}$$

where γ is a number such that $0 \leq \gamma < 1$, ϵ is any positive number, and K is a constant dependent only on ϵ and x_0 .

Then the function $f(z)$ defined by (2.1), *when considered for values of z of large modulus lying within the fixed sector $\gamma\pi < \arg z < (2 - \gamma)\pi$, is developable asymptotically in the form*

$$f(z) \sim -\sum_{n=1}^{\infty} g(-n)z^{-n}. \tag{2.3}$$

Moreover, in case conditions (a) *and* (b) *are satisfied except that $g(w)$ has a singularity situated at the point $w = w_1$, where w_1 is not a negative integer, then* (2.3) *continues to hold, provided the loop integral*

$$\frac{1}{2i}\int_C \frac{g(w)(-z)^w}{\sin \pi w}\, dw \tag{2.4}$$

is subtracted from the right number. The loop C surrounds the point w_1 and extends to infinity in any convenient direction. (*For full description of the loop, see* [2; 9].)

THEOREM II. *Let condition* (a) *of Theorem* I *remain as stated, while condition* (b) *is replaced by a new condition* (b′) *wherein* (2.2) *is changed to read*

$$| g(x + iy) | < Ke^{(k\pi+\epsilon)|y|}, \tag{2.5}$$

where k is a positive odd integer. Then the function $f(z)$, when considered for values of z such that $| \arg z | < \pi$, may be expressed in the form

$$f(z) = \int_{-l-\frac{1}{2}}^{\infty} g(x)z^x \frac{\sin k\pi x}{\sin \pi x}\, dx - \sum_{n=1}^{l} g(-n)z^{-n} + \xi(z, l), \tag{2.6}$$

where l is any positive integer, and where the function $\xi(z, l)$ satisfies

(2.7) $$\lim_{|z|\to\infty} z^l \xi(z, l) = 0.$$

We note that Theorem II does not furnish complete information on the asymptotic development of $f(z)$, because of the presence of the integral in the right member of (2.6). We may also remark that if k is an even integer, then the expression z^w occurring in (2.6) becomes $(-z)^w$. In the applications which we make of Theorem II, k is always an odd integer.

3. **Extension of Theorem** II. It can be shown that if conditions (a) and (b′) of Theorem II are satisfied except that $g(w)$ has a singularity situated at the point $w = w_1$, where w_1 is not a negative integer, then the theorem continues to hold provided one subtracts from the right member of (2.6) the loop integral

(3.1) $$\frac{1}{2i}\int_C \frac{g(w)z^w}{e^{k\pi i w}\sin \pi w}\,dw,$$

where the loop C surrounds the point w_1 and extends to infinity in any convenient direction lying in either the third or the fourth quadrant. If k is an even integer, then the expression z^w appearing in (3.1) is to be changed to $(-z)^w$. Ford [2; 36] has proved this assertion for the special case in which k is unity. The proof of the general case, recently obtained by one of the present authors, is expected to appear in detail elsewhere and will not be given here.

II. The Function $F_\alpha(z, t)$

4. **Properties of the coefficient.** As was stated in §1, Part II of this paper will be devoted to the determination of the asymptotic developments of the general function

(4.1) $$F_\alpha(z, t) = \sum_{n=0}^{\infty} \frac{h(n)}{\Gamma(\alpha n + t)}\, z^n,$$

where t is any constant, real or complex, and $\alpha > 0$. We shall now impose on the function $h(n)$ such conditions as will make the two fundamental theorems stated in §2 applicable to the series appearing in (4.1). Let $h(n)$ be considered as a function $h(w)$ of $w = x + iy$, and let us suppose that, when considered throughout any arbitrary right half-plane $x > x_0$, it satisfies the following two conditions:

(A) the function $g(w) = \dfrac{h(w)}{\Gamma(\alpha w + t)}$ is single-valued and analytic;

(B) the function $h(w)$, when considered for values of w of large modulus, can be expressed in the form

$$(4.2)\quad h(w) = c_0 + \frac{c_1}{\alpha w + t} + \cdots + \frac{c_s + \delta(\alpha w, s)}{(\alpha w + t)(\alpha w + t + 1) \cdots (\alpha w + t + s)},$$

where c_0, c_1, $\cdots$ are independent of w, and where $\delta(\alpha w, s)$ is such that $\lim_{|w| \to \infty} \delta(\alpha w, s) = 0$, $s = 1, 2, 3, \cdots$.

It then follows that for all $x > x_0$ and for all $|y|$ sufficiently large we may write

$$(4.3)\qquad \left|\frac{h(w)}{\Gamma(\alpha w + t)}\right| < Ke^{(\alpha\pi/2+\epsilon)|y|},$$

where ϵ is any positive number chosen in advance, and K is a constant dependent only on ϵ and x_0. For the property of the gamma function used here, see [2; 61]. Hence Theorem I of §2 is applicable to the series (4.1) whenever $0 < \alpha < 2$. Moreover, Theorem II is applicable when α has any positive value, k being suitably chosen as indicated in the next section.

5. **The improper integral.** If k is selected as the smallest odd integer such that $2k > \alpha$, then, upon applying Theorem II of §2 to the function $F_\alpha(z, t)$, we obtain the relation

$$(5.1)\qquad F_\alpha(z, t) = \int_{-l-\frac{1}{2}}^{\infty} \frac{h(x)z^x \sin k\pi x}{\Gamma(\alpha x + t) \sin \pi x}\, dx - \sum_{m=1}^{l} \frac{h(-m)z^{-m}}{\Gamma(t - \alpha m)} + \xi(z, l),$$

where l and $\xi(z, l)$ are described in the statement of the theorem. We now undertake to determine the asymptotic behavior for large $|z|$ of the integral appearing in the right member of (5.1). In what follows, the expression $z^{1/\alpha} \times e^{2\pi i\mu/\alpha}$ is prominently involved. We shall denote it by Z_μ; thus e^{Z_μ} will mean $\exp(z^{1/\alpha}e^{2\pi i\mu/\alpha})$, etc. For our present purpose, we shall establish the following lemma.

LEMMA. *Let the function $h(w)$ satisfy conditions* (A) *and* (B) *of* §4. *Then, if $\alpha > 0$ and l' is a sufficiently large positive number, we have*

$$(5.2)\qquad \int_{-l'}^{\infty} \frac{h\left(\dfrac{x}{\alpha}\right)z^x}{\Gamma(x + t)}\, dx \sim \begin{cases} e^z z^{1-t} \displaystyle\sum_{n=0}^{\infty} c_n z^{-n}; & |\arg z| \le \frac{1}{2}\pi, \\ 0; & \frac{1}{2}\pi < |\arg z| < \pi, \end{cases}$$

where the c_i are those appearing in (4.2).

The conclusion for $|\arg z| \le \frac{1}{2}\pi$ has already been established by Ford. In order to show that the integral vanishes for $|\arg z| > \frac{1}{2}\pi$, when $|z| \to \infty$, as indicated by the second conclusion, we first note that the function $h(x/\alpha)$ is

bounded along the path of integration. Consequently, the integral in question may, for the purpose, be compared with the expression

$$z^{1-t}\int_{-l''}^{\infty}\frac{z^x\,dx}{\Gamma(x+1)};\qquad l''=l'+1-t.$$

But if L is the greatest integer in the real part of l'', we have

$$z^{1-t}\int_{-l''}^{\infty}\frac{z^x\,dx}{\Gamma(x+1)}=z^{1-t}[e^z+\eta(z,L)];\qquad |\arg z|<\pi,$$

where $\lim_{|z|\to\infty} z^L\eta(z,L)=0$. For proof of this statement, see [2; 65, (10) et seq.]. Also for proof of the first conclusion of the lemma, see [2; 67, (19)]. The function $h(x)$ there appearing is the same as that denoted by $h(x/\alpha)$ in (5.2) above. Since e^z vanishes as $|z|\to\infty$ whenever the real part of z is negative, the conclusion stated in the lemma for $|\arg z|>\frac{1}{2}\pi$ clearly follows.

Having established the preceding lemma, let us now, in (5.2), replace x by αx, l' by $l+\frac{1}{2}$, and z by Z_μ. Then we obtain the relation

$$\text{(5.3)}\qquad \int_{-l-\frac{1}{2}}^{\infty}\frac{h(x)z^x}{\Gamma(\alpha x+t)}e^{2\pi i\mu x}\,dx\sim\begin{cases}\dfrac{1}{\alpha}e^{Z_\mu}(Z_\mu)^{1-t}\displaystyle\sum_{0}^{\infty}c_n(Z_\mu)^{-n},\\[2ex] 0,\end{cases}$$

the first result holding when $|\arg z+2\pi\mu|\leq\frac{1}{2}\pi\alpha$, and the second otherwise.

We are now able to write the asymptotic development of the integral in question, namely, that appearing in the right member of (5.1). For, upon taking account of the identity

$$\frac{\sin k\pi x}{\sin\pi x}=\sum_{\mu=-p}^{p}e^{2\pi i\mu x},\qquad 2p+1=k,$$

and applying (5.3) to the separate terms, we have

$$\text{(5.4)}\qquad \int_{-l-\frac{1}{2}}^{\infty}\frac{h(x)z^x\sin k\pi x}{\Gamma(\alpha x+t)\sin\pi x}\,dx\sim\frac{1}{\alpha}\sum_{\mu}e^{Z_\mu}(Z_\mu)^{1-t}\Big\{\sum_{n=0}^{\infty}c_n(Z_\mu)^{-n}\Big\},$$

where the symbol $\sum_{\mu}$ denotes summation taken over those integral values of μ for which $|\arg z+2\pi\mu|\leq\frac{1}{2}\pi\alpha$.

When $0<\alpha<2$, the value of k is unity, and the sum in (5.4) consists of a single term, that one in which $\mu=0$. Thus when $|\arg z|\leq\frac{1}{2}\pi\alpha$, we have

$$\text{(5.5)}\qquad \int_{-l-\frac{1}{2}}^{\infty}\frac{h(x)z^x}{\Gamma(\alpha x+t)}\,dx\sim\frac{1}{\alpha}z^{(1-t)/\alpha}\exp z^{1/\alpha}\Big\{\sum_{n=0}^{\infty}c_n z^{-n/\alpha}\Big\}.$$

6. **Asymptotic developments of $F_\alpha(z,t)$.** The results in the previous section having been established, we shall now obtain the asymptotic developments of the function

$$F_\alpha(z, t) = \sum_{n=0}^{\infty} \frac{h(n)z^n}{\Gamma(\alpha n + t)} \qquad (\alpha > 0), \tag{6.1}$$

where t is any constant, real or complex, and where the function $h(n)$ satisfies the conditions named in §4. We prove the following general theorem.

THEOREM. *Consider the function $F_\alpha(z, t)$ defined by* (6.1). *If the function $h(n)$, when considered as a function $h(w)$ of $w = x + iy$, satisfies conditions* (A) *and* (B) *described in §4, then $F_\alpha(z, t)$ has the following asymptotic developments*:

(i) $\quad 0 < \alpha < 2; \quad \frac{1}{2}\pi\alpha < \arg z < (2 - \frac{1}{2}\alpha)\pi.$

$$F_\alpha(z, t) \sim -\sum_{n=1}^{\infty} \frac{h(-n)}{\Gamma(t - \alpha n)} z^{-n}. \tag{6.2}$$

(ii) $\quad 0 < \alpha < 2; \quad |\arg z| < \frac{1}{2}\pi\alpha.$

$$F_\alpha(z, t) \sim \frac{1}{\alpha} z^{(1-t)/\alpha} \exp(z^{1/\alpha}) \sum_{n=0}^{\infty} c_n z^{-n/\alpha}, \tag{6.3}$$

where the c_0, c_1, $\cdots$ are those appearing in (4.2).

(iii) $\quad 0 < \alpha < 2; \quad |\arg z| = \frac{1}{2}\pi\alpha.$

$$F_\alpha(z, t) \sim \frac{1}{\alpha} z^{(1-t)/\alpha} \exp(z^{1/\alpha}) \sum_{n=0}^{\infty} c_n z^{-n/\alpha} - \sum_{n=1}^{\infty} \frac{h(-n)}{\Gamma(t - \alpha n)} z^{-n}. \tag{6.4}$$

(iv) $\quad \alpha \geq 2; \quad |\arg z| < \pi.$

$$F_\alpha(z, t) \sim \frac{1}{\alpha} \sum_{\mu} \{e^{Z_\mu} Z_\mu^{(1-t)} \sum_{n=0}^{\infty} c_n (Z_\mu)^{-n}\}, \tag{6.5}$$

the first summation being taken over those integral values of μ for which

$$|\arg z + 2\pi\mu| \leq \tfrac{1}{2}\pi\alpha.$$

In order to establish (6.2), we note that the function

$$g(w) = \frac{h(w)}{\Gamma(\alpha w + t)}$$

satisfies conditions (a) and (b) of Theorem I, §2, where $\gamma = \frac{1}{2}\pi\alpha$. If we apply this theorem, (6.2) follows at once.

To obtain (6.3) and (6.4), we first note that since $0 < \alpha < 2$ and $k = 1$, equation (5.1) takes the form

$$F_\alpha(z, t) = \int_{-l-\frac{1}{2}}^{\infty} \frac{h(x)z^x}{\Gamma(\alpha x + t)}\, dx - \sum_{m=1}^{l} \frac{h(-m)z^{-m}}{\Gamma(t - \alpha m)} + \xi(z, l), \tag{6.6}$$

where we have $\lim_{|z| \to \infty} z^l \xi(z, l) = 0$, and $|\arg z| < \pi$. If we replace the integral appearing in the right member of (6.6) by its asymptotic development as given by (5.5) and write the resulting relation in asymptotic notation, we have (6.4).

To obtain (6.3), we have only to factor the expression $\exp z^{1/\alpha}$ out of the right member of (6.4) and then confine $\arg z$ to the narrower range $|\arg z| < \frac{1}{2}\pi\alpha$. For since $\exp(-z^{1/\alpha})$ approaches zero as $|z| \to \infty$, when $\arg z$ is so restricted, (6.4) immediately reduces to (6.3). A similar argument employing (5.4) will lead to the result given in (6.5).

7. **The function** $E_\alpha(z)$. It is worthy of note that, if in the theorem of the previous section we set $h(w) = t = 1$, we obtain immediately the asymptotic developments of the special function

$$E_\alpha(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(\alpha n + 1)}.$$

These well-known results were first established by Mittag-Leffler [3].

8. **Remarks and generalizations.** It is desirable to add some supplementary observations relative to the general theorem which was proved in §6.

We note that in order to secure simplicity of statement, condition (A) of §4, relative to the coefficient $g(n)$, was made unnecessarily restrictive. In case condition (A) is not satisfied but, instead, the function

$$g(w) = \frac{h(w)}{\Gamma(\alpha w + t)}, \tag{8.1}$$

while still remaining single-valued throughout the halfplane $x > x_0$, has a singularity situated at the point $w = w_1$, which is not a negative integer, then the theorem continues to hold provided one subtracts from the right member of (6.2), (6.3), (6.4), (6.5) the appropriate loop integral, as called for by Theorem I, §2, or by the extension of Theorem II obtained in §3. If the singular point w_1 is a negative integer, say $-m$, then in (5.1) the term $-g(-m)z^{-m}$ appearing in the right member is to be suppressed, and in its place is to be supplied the proper loop integral about the singular point. Finally, whenever the singular point w_1 is polar in character, the corresponding loop integral is equivalent to the residue of the function

$$\frac{\pi g(w) z^w}{e^{k\pi i w} \sin \pi w} \tag{8.2}$$

at w_1, where in (8.2) k is to be assigned its proper value. The extension of the above remarks to the case in which more than one singular point is present is obvious.

As a second restriction, we have kept α real and positive. The methods which we have employed are, however, still applicable in case α assumes complex values having positive real part.

As was mentioned in §1, the function $F_\alpha(z, t)$ has been studied recently by Wright. The conditions which he has imposed on the function (8.1) are equiv-

alent to conditions (A) and (B) of §4 above. The asymptotic developments (6.1)–(6.5) are obtained, though they are expressed in a briefer form. The case in which the function (8.1) has a finite number of poles is included, but no mention is made of the case in which a singular point is non-polar. Thus the present paper, based on the theorems of Ford and Newsom, obtains all the results obtained by Wright, together with the additional point just mentioned.

III. The Function $E_\alpha(z, \theta, \beta)$

9. **Properties of the coefficient.** As an application of the theorem obtained in Part II, together with the remarks made in the first paragraph of §8, we now find the asymptotic developments of the special function denoted by Barnes as $E_\alpha(z, \theta, \beta)$, namely,

$$E_\alpha(z, \theta, \beta) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(\alpha n + 1)(n + \theta)^\beta}, \tag{9.1}$$

where β and θ are any constants, real or complex, except that θ is not a negative integer. Evidently we may write

$$E_\alpha(z, \theta, \beta) = \sum_{n=0}^{\infty} \frac{h(n) z^n}{\Gamma(\alpha n + \beta + 1)}, \tag{9.2}$$

where

$$h(n) = \frac{\Gamma(\alpha n + \beta + 1)}{\Gamma(\alpha n + 1)(n + \theta)^\beta}. \tag{9.3}$$

The coefficient of z^n in (9.2), when considered as a function of $w = x + iy$, is single-valued and analytic throughout the finite w plane except for the singularity at the point $w = -\theta$. Moreover, it can be shown that $h(w)$, when considered for values of w of large modulus lying in the arbitrary right halfplane $x > x_0$, can be expressed in the form

$$\begin{aligned} h(w) = \alpha^\beta \Big\{ 1 &+ \frac{q_1}{\alpha w + \beta + 1} + \frac{q_2}{(\alpha w + \beta + 1)(\alpha w + \beta + 2)} + \cdots \\ &+ \frac{q_n + \delta(\alpha w, n)}{(\alpha w + \beta + 1)(\alpha w + \beta + 2) \cdots (\alpha w + \beta + n)} \Big\}, \end{aligned} \tag{9.4}$$

where the quantities $q_1, q_2, \cdots$ are independent of w, and where $\lim_{|w| \to \infty} \delta(\alpha w, n) = 0$. The proof of this fact follows with slight modifications (due to the presence of α) the proof of a similar statement appearing in the work of Ford [2; 69]. As for the values of the q's, they may be calculated as follows: Let

$$R_n(w) = \alpha^{-\beta} h(w) - 1 - \sum_{s=0}^{n-1} \frac{q_{s+1}}{(\alpha w + \beta + 1) \cdots (\alpha w + \beta + s + 1)}.$$

Then we have

$$q_{n+1} = \lim_{|w|\to\infty} [(\alpha w)^{n+1} R_n(w)]. \tag{9.5}$$

10. **The loop integral.** Inasmuch as the function

$$g(w) = \frac{h(w)}{\Gamma(\alpha w + \beta + 1)}$$

satisfies conditions (A) and (B) of §4, we may apply the general theorem of §6 to the function $E_\alpha(z, \theta, \beta)$. Because of the singularity at $w = -\theta$, there must be subtracted from the right members of (6.2)-(6.4) the loop integral

$$I_\alpha(z, \theta, \beta) = \frac{1}{2i} \int_c \frac{z^w\, dw}{\Gamma(\alpha w + 1)(w + \theta)^\beta \sin \pi w}. \tag{10.1}$$

In the case of (6.5), the integral to be subtracted has the additional factor $e^{k\pi i w}$ appearing in the denominator of the integrand. In order to determine the complete asymptotic development of $E_\alpha(z, \theta, \beta)$, we must first find such a development for the above loop integral.

Let us agree to take the loop C as extending to infinity in a direction of the third quadrant. Then, if we make the transformation $w' = -(w + \theta)$ and then drop the primes, the two integrals under consideration take the forms

$$I_\alpha(z, \theta, \beta) = \frac{(-z)^{-\theta}}{2i} \int_\gamma \frac{(-z)^w(-w)^{-\beta} dw}{\Gamma(1 - \alpha w - \alpha\theta) \sin \pi(w + \theta)} \qquad (\alpha < 2), \tag{10.2}$$

$$I_\alpha(z, \theta, \beta) = \frac{(-z)^{-\theta}}{2i} \int_\gamma \frac{(-z)^w(-w)^{-\beta} e^{(k-1)\pi i(w+\theta)}}{\Gamma(1 - \alpha w - \alpha\theta) \sin \pi(w + \theta)}\, dw \qquad (\alpha \geq 2). \tag{10.3}$$

The new loop γ now extends to infinity in a direction of the first quadrant.

Now if θ is not an integer, or if both θ and $\alpha\theta$ are integers, then the following series developments exist and are convergent in the neighborhood of $w = 0$:

$$\frac{1}{\Gamma(1 - \alpha w - \alpha\theta) \sin \pi(w + \theta)} = \sum_{n=0}^{\infty} a_n(-w)^n \qquad (\alpha < 2); \tag{10.4}$$

$$\frac{e^{(k-1)\pi i(w+\theta)}}{\Gamma(1 - \alpha w - \alpha\theta) \sin \pi(w + \theta)} = \sum_{n=0}^{\infty} b_n(-w)^n \qquad (\alpha \geq 2). \tag{10.5}$$

Under these conditions, the asymptotic development of $I_\alpha(z, \theta, \beta)$ is given by a theorem due to Barnes, and included, together with numerous applications, in the work of Ford. For the statement and proof of the theorem, see [2; 16–22, II]. According to this theorem, we have

$$I_\alpha(z, \theta, \beta) \sim -(-z)^{-\theta} \sum_{n=0}^{\infty} \frac{A_n}{[\log(-z)]^{1-\beta+n}\, \Gamma(\beta - n)}, \tag{10.6}$$

where A_n is equal to a_n or b_n according as $\alpha < 2$ or $\alpha \geq 2$, and where $\log(-z) = \log|z| + i(\phi - \pi)$; $0 \leq \phi < 2\pi$. For the definitions of a_n and b_n see (10.4) and (10.5). Thus, (10.6) furnishes the required development under the conditions stated concerning θ and $\alpha\theta$.

If θ is an integer and $\alpha\theta$ is not, then developments (10.4) and (10.5) are not possible, and (10.6) does not hold. However, when θ is an integer, $\sin \pi(w + \theta)$ reduces to $(-1)^\theta \sin \pi w$, and (10.2) may be rewritten in the special form

$$I_\alpha(z, \theta, \beta) = \frac{z^{-\theta}}{2\pi i}\int_\gamma \frac{(-z)^{-w}(-w)^{-\beta-1}\,\pi w}{\sin \pi w\;\Gamma(1 - \alpha w - \theta)}\,dw. \tag{10.7}$$

Moreover, (10.3) may be rewritten in a similar manner. Now there do exist power series, convergent in the neighborhood of $w = 0$, such that

$$\frac{\pi w}{\Gamma(1 - \alpha w - \alpha\theta)\sin \pi w} = \sum_{n=0}^{\infty} d_n(-w)^n \qquad (\alpha < 2), \tag{10.8}$$

$$\frac{\pi w e^{(k-1)\pi w}}{\Gamma(1 - \alpha w - \alpha\theta)\sin \pi w} = \sum_{n=0}^{\infty} e_n(-w)^n \qquad (\alpha \geq 2). \tag{10.9}$$

Hence we may again apply the theorem of Barnes and write

$$I_\alpha(z, \theta, \beta) \sim -z^{-\theta}\sum_{n=0}^{\infty}\frac{D_n}{[\log(-z)]^{n-\beta}\,\Gamma(1 + \beta - n)}, \tag{10.10}$$

in which D_n is equal to d_n or e_n according as $\alpha < 2$ or $\alpha \geq 2$, and where $\log(-z)$ is to be interpreted as in (10.6).

It is to be noted that, if β is an integer, then both (10.6) and (10.10) reduce to finite series, since the function $1/\Gamma(\beta - n)$ vanishes when $n \geq \beta$. The loop integral $I_\alpha(z, \theta, \beta)$, of course, reduces to a residue in such a case.

11. **Asymptotic developments of $E_\alpha(z, \theta, \beta)$.** We shall now obtain the asymptotic developments of the function $E_\alpha(z, \theta, \beta)$ defined by (9.1). The developments will involve the constants q_n ($n = 1, 2, 3, \cdots$) appearing in (9.4), as well as the constants A_n and D_n appearing in (10.6) and (10.10), respectively. Moreover, the specific form of some of the terms depends on whether or not θ is an integer. We shall establish the following

THEOREM. *If θ is not an integer, or if both θ and $\alpha\theta$ are integers, then, for values of z of large modulus, the function $E_\alpha(z, \theta, \beta)$ defined by* (9.1) *has the following asymptotic developments*:

(i) $\quad 0 < \alpha < 2; \quad \frac{1}{2}\alpha\pi < \arg z < (2 - \frac{1}{2}\alpha)\pi.$

$$E_\alpha(z, \theta, \beta) \sim \sum_{n=1}^{\infty}\frac{z^{-n}}{\Gamma(1 - \alpha n)(\theta - n)^\beta} + (-z)^{-\theta}[\log(-z)]^{\beta-1}\sum_{n=0}^{\infty}\frac{A_n}{[\log(-z)]^n\,\Gamma(\beta - n)}. \tag{11.1}$$

(ii) $$0 < \alpha < 2; \qquad |\arg z| < \tfrac{1}{2}\alpha\pi.$$

$$E_\alpha(z, \theta, \beta) \sim \alpha^{\beta-1} z^{-\beta/\alpha} \exp z^{1/\alpha} \sum_{n=0}^{\infty} q_n z^{-n/\alpha}. \tag{11.2}$$

(iii) $$0 < \alpha < 2; \qquad |\arg z| = \tfrac{1}{2}\alpha\pi.$$

$$\begin{aligned} E_\alpha(z, \theta, \beta) \sim & \sum_{n=1}^{\infty} \frac{z^{-n}}{\Gamma(1 - \alpha n)(\theta - n)^\beta} \\ & + \alpha^{\beta-1} z^{-\beta/\alpha} \exp z^{1/\alpha} \sum_{n=0}^{\infty} q_n z^{-n/\alpha} \\ & + (-z)^{-\theta} [\log(-z)]^{\beta-1} \sum_{n=0}^{\infty} \frac{A_n}{[\log(-z)]^n \, \Gamma(\beta - n)}. \end{aligned} \tag{11.3}$$

(iv) $$\alpha \geq 2; \qquad |\arg z| < \pi.$$

$$E_\alpha(z, \theta, \beta) \sim \alpha^{\beta-1} \sum_{\mu} \{e^{Z_\mu} Z_\mu^{-\beta} \sum_{n=0}^{\infty} q_n Z_\mu^{-n}\}, \tag{11.4}$$

where the first summation is taken over those integral values of μ *which satisfy the inequality* $|\arg z + 2\pi\mu| \leq \frac{1}{2}\alpha\pi$.

Moreover, if θ *is an integer while* $\alpha\theta$ *is not an integer, then* (11.1) *becomes*

$$\begin{aligned} E_\alpha(z, \theta, \beta) \sim & {\sum_{n=1}^{\infty}}' \frac{z^{-n}}{\Gamma(1 - \alpha n)(\theta - n)^\beta} \\ & + z^{-\theta} [\log(-z)]^{\beta} \sum_{n=0}^{\infty} \frac{D_n}{[\log(-z)]^n \, \Gamma(1 + \beta - n)}, \end{aligned} \tag{11.5}$$

while (11.3) *becomes*

$$\begin{aligned} E_\alpha(z, \theta, \beta) \sim & {\sum_{n=1}^{\infty}}' \frac{z^{-n}}{\Gamma(1 - \alpha n)(\theta - n)^\beta} \\ & + \alpha^{\beta-1} q^{-\beta/\alpha} \exp z^{1/\alpha} \sum_{n=0}^{\infty} q_n z^{-n/\alpha} \\ & + z^{-\theta} [\log(-z)]^{\beta} \sum_{n=0}^{\infty} \frac{D_n}{[\log(-z)]^n \, \Gamma(1 + \beta - n)}, \end{aligned} \tag{11.6}$$

where $\sum'$ *means that the term* $n = \theta$ *is deleted from the sum. Developments* (11.2) *and* (11.3) *remain unaltered. In all the determinations, if* $z = re^{i\phi}$, *then* $\log(-z) = \log r + i(\phi - \pi)$, *where* $0 \leq \phi < 2\pi$.

In fact, the above asymptotic developments are obtained from (6.1)–(6.5) upon replacing t appearing there by $\beta + 1$, c_n by $\alpha^\beta q_n$, and then subtracting from the right-hand members the appropriate development of the loop integral $I_\alpha(z, \theta, \beta)$

as obtained in §10. In the case of (11.2), (11.3) and (11.5), the final form is obtained by factoring the expression $\exp(z^{1/\alpha})$ out of the algebraic and logarithmic terms, and noting that $\exp(-z^{1/\alpha})$ approaches zero as $z \to \infty$.

Developments (11.1)–(11.4) were first obtained by Barnes, and appear in [1]. However, developments (11.5) and (11.6) do not appear. They could doubtlessly be obtained by a proper extension of Barnes' analysis.

Bibliography

1. E. W. Barnes, *The asymptotic expansions of functions defined by Taylor's series*, Philosophical Transactions of the Royal Society of London, 206(A), (1906), pp. 249–297.
2. W. B. Ford, *The asymptotic developments of functions defined by Maclaurin series*, University of Michigan Studies, Scientific Series, vol. XI(1936).
3. G. Mittag-Leffler, *Sur la nouvelle fonction* ⋯ , Comptes Rendus, vol. 137(1903), p. 554.
4. C. V. Newsom, *On the character of certain entire functions in distant portions of the plane*, American Journal of Mathematics, vol. 60(1938), pp. 561–572.
5. E. M. Wright, *The asymptotic expansion of integral functions defined by Taylor series*, Philosophical Transactions of the Royal Society of London, series A, vol. 238(1940), pp. 423–451.
6. E. M. Wright, *The asymptotic expansion of the generalized hypergeometric function*, Proceedings of the London Mathematical Society, (2), vol. 46(1940), pp. 389–408.
7. E. M. Wright, *The generalized Bessel function of order greater than one*, Quarterly Journal of Mathematics, Oxford Series, vol. 11(1940), pp. 36–48.

Purdue University.

THE ABSOLUTE CONVERGENCE OF FOURIER SERIES

By Min-Teh Cheng

1. **Introduction.** In this note we suppose throughout that the functions f_1, f_2 and f are integrable and periodic with period 2π. $f(x)$ is said to be a Young's continuous function [9], if there exist two functions f_1 and f_2 of the Lebesgue class $L^2(-\pi, \pi)$, satisfying

$$f(x) = \frac{1}{\pi} \int_{-\pi}^{\pi} f_1(\xi) f_2(\xi + x)\, d\xi.$$

The necessary and sufficient condition for the absolute convergence of a trigonometric series in the whole interval is that the series be a Fourier series of a Young's continuous function [3], [4]. One object of this paper is to obtain Young's functions with conditions imposing more on f_1 and less on f_2. Indeed, we shall prove

Theorem 1. *If $f_1 \in \mathrm{Lip}\,(\alpha, p)$, $\alpha p > \frac{1}{2}$, $2 \geq p > 1$, $\alpha\alpha \leq 1$ and $f_2 \in \mathrm{Lip}\,(1/2p, q)$ for $q > 1$, then the Fourier series*

$$f(x) = \frac{1}{\pi} \int_{-\pi}^{\pi} f_1(\xi) f_2(\xi + x)\, d\xi$$

converges absolutely in the whole interval.

The notation $\varphi \in \mathrm{Lip}\,(\alpha, p)$ means that

$$\left(\int_{-\pi}^{\pi} |\, \Delta\varphi \,|^p \, d\theta \right)^{1/p} = O(h^\alpha) \tag{1.1}$$

as $h \to +0$ and $\Delta\varphi$ denotes one of the three differences [5]

$$\varphi(\theta) - \varphi(\theta - h), \qquad \varphi(\theta + h) - \varphi(\theta), \qquad \varphi(\theta + h) - \varphi(\theta - h).$$

If $f(\theta) \in \mathrm{Lip}\,(\alpha, p)$ and

$$f(\theta) \sim \sum_{m=-\infty}^{\infty} c_m e^{im\theta}, \tag{1.2}$$

then the series

$$\sum_{m=-\infty}^{\infty} |\, c_m \,|^k \tag{1.3}$$

is convergent for $k > p/(p + \alpha p - 1)$, if

$$0 < \alpha \leq 1, \qquad 1 < p \leq 2.$$

Received April 15, 1942.

This is known as the theorem of Titchmarsh [8] and is a generalization of a theorem of Szász [7]. The original result in this direction is due to Bernstein [1]; this result shows that the Fourier series of a function of Lip α ($> \frac{1}{2}$) is absolutely convergent. Various extensions are known [2], [6].

If $\alpha p > 1$, then $p/(p + \alpha p - 1) < 1$. Therefore, Titchmarsh's result involves absolute convergence in this case. We can, however, establish

THEOREM 2. *If $1 < p \leq 2$, $\epsilon > 0$, and, as $h \to +0$,*

$$\int_{-\pi}^{\pi} |\Delta f|^p \, d\theta = O\{h(\log h^{-1})^{-p-\epsilon}\}, \tag{1.4}$$

then the Fourier series of $f(\theta)$ is absolutely convergent. Moreover, the positive number ϵ in (1.4) cannot be replaced by 0.

The first part of this proposition is contained in

THEOREM 3. *If $\alpha > 0$, $1 < p \leq 2$, and, as $h \to +0$,*

$$\begin{aligned}&\int_{-\pi}^{\pi} |\Delta f|^p \, d\theta \\ &\quad = O(h(\log_1 h^{-1})^{-p}(\log_2 h^{-1})^{-p} \cdots (\log_{k-1} h^{-1})^{-p}(\log_k h^{-1})^{-(1+\alpha p)}),\end{aligned} \tag{1.5}$$

then, for $T < \alpha + p^{-1} - 1$, the series

$$\sum_m{}'' \, | c_m | (\log_k | m |)^T \tag{1.6}$$

is convergent.

The notation $\log_n x$ means $\log (\log_{n-1} x)$ and $\log_1 x = \log x$. $\sum''$ denotes the summation for m, where the terms containing

$$\log_{n-1} | m | \not> 0 \qquad (n = 1, 2, \cdots, k)$$

are omitted. If $k = 1$, we have

THEOREM 4. *If $\alpha > 0$, $1 < p \leq 2$, and, as $h \to +0$,*

$$\int_{-\pi}^{\pi} |\Delta f|^p \, d\theta = O(h(\log h^{-1})^{-1-\alpha p}), \tag{1.7}$$

then, for $T < \alpha + p^{-1} - 1$, the series

$$\sum_{-\infty}^{\infty}{}' \, | c_m | (\log | m |)^T \tag{1.8}$$

converges. ($\sum'$ denotes summation for m, with the term corresponding to $m = 0$ omitted.) Further, the series

$$\sum_{-\infty}^{\infty}{}' \, | c_m | (\log | m |)^{T_0}$$

with $T_0 = \alpha + p^{-1} - 1$ may be divergent.

The following theorem is an immediate consequence of Theorem 2.

THEOREM 5. *If $f(x) \varepsilon$ Lip (α, p), $p < 2$, $\alpha p = 1$, and*

$$\Delta f = O((\log h^{-1})^{-b}) \qquad (0 \leq x \leq 2\pi), \tag{1.9}$$

where $b > 2/(2 - p)$, then the Fourier series $f(x)$ converges absolutely.

In fact, let $p_2 > 2 > p$; then Hölder's inequality gives

$$\int_0^{2\pi} |\Delta f|^2\, dx \leq \left\{\int_0^{2\pi} |\Delta f|^p\, dx\right\}^{(p_2-2)/(p_2-p)} \left\{\int_0^{2\pi} |\Delta f|^{p_2}\, dx\right\}^{(2-p)/(p_2-p)}. \tag{1.10}$$

Setting $p_2 = \infty$, we have

$$\int_0^{2\pi} |\Delta f|^2\, dx = O(h(\log h^{-1})^{-b(2-p)}). \tag{1.11}$$

The required conclusion follows from Theorem 2, since $b(2 - p) > 2$.

The functions of the class Lip (1, 1) are of bounded variation. Zygmund [10] proves that "if $f(x)$ is of bounded variation and satisfies

$$|\Delta f| \leq c \log^{-2-\eta} h^{-1} \qquad (\eta > 0, 0 \leq x \leq 2\pi),$$

then the Fourier series $f(x)$ converges absolutely". This is obtained by putting $\alpha = p = 1$ in Theorem 5. The example $f(x) = \sum (\sin nx)/(n \log n)$ shows that the theorem fails if $\eta = -1$. But the question whether the theorem is valid for the case $0 \geq \eta > -1$ remains open.

The arguments of the present paper are based on the inequalities of Hausdorff, by means of which we obtain incidentally a simple proof of Titchmarsh's theorem.

2. **A simple proof of Titchmarsh's theorem.** Let $1 < p \leq 2$, $1/p + 1/p' = 1$. The theorem of Hausdorff states that, if the series $\sum |c_m|^p$ is convergent, $f(\theta)$ belongs to $L^{p'}$, and that, if $f(\theta)$ belongs to L^p, the series $\sum |c_m|^{p'}$ is convergent. The corresponding theorem for the functions of Lip (α, p) is required for our proof of Titchmarsh's theorem, and may be stated as follows.

LEMMA 1. *Let $1 < p \leq 2$, $p^{-1} + p'^{-1} = 1$, $0 < \alpha \leq 1$. If the series*

$$\sum |c_m m^\alpha|^p$$

is convergent, then $f \varepsilon$ Lip (α, p'). If $f \varepsilon$ Lip (α, p), then $\sum |c_m m^{\alpha'}|^{p'}$ $(\alpha' < \alpha)$ converges, but the series

$$\sum |c_m m^\alpha|^{p'}$$

may be divergent.

The last clause of Lemma 1 can be verified by the function [5; 632]

$$f(\theta) = |\theta|^{-a}, \qquad (a + 1)p > 1.$$

This function belongs to Lip (α, p) with $\alpha = 1/p - a$, but the Fourier constants are of the exact order $| m |^{a-1}$. Accordingly, the series

$$\sum | c_m m^\alpha |^{p'} > A \sum (| m |^{a-1+\alpha})^{p'} = A \sum | m |^{-1}$$

diverges.

The convergence of $\sum | c_m m^\alpha |^p$ implies

$$\gamma_n = \sum_{-n}^{n} | c_m m |^p = O(n^{p-\alpha p}). \tag{2.1}$$

From

$$\Delta f(\theta) = f(\theta + h) - f(\theta - h) \sim 2i \sum_{-\infty}^{\infty} c_m e^{im\theta} \sin mh,$$

it follows, by (2.1), that

$$\begin{aligned}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} | \Delta f |^{p'} d\theta\right)^{p/p'} &\leq 2^p(\sum_{-\infty}^{\infty} | c_m \sin mh |^p)\\ &\leq 2^p[h^p \sum_{h|m|\leq 1} | c_m m |^p + \sum_{1<h|m|} | c_m |^p]\\ &\leq A[h^p\gamma_{[h^{-1}]} + \sum_{1<hm} \gamma_m(m^{-p} - (m+1)^{-p}) = O(h^{\alpha p});\end{aligned}$$

hence $f(\theta)$ belongs to Lip (α, p').

To prove the second part of the lemma, write

$$\Phi(h) = \left(\frac{1}{2\pi}\int_0^{2\pi} | \Delta f |^p d\theta\right)^{p'/p}; \tag{2.2}$$

then $\Phi(h) = O(h^{\alpha p'})$, by hypothesis. Hence the integral

$$\int_0^{2\pi} \frac{\Phi(h)}{h^{1+\alpha p'-\epsilon}} dh \tag{2.3}$$

exists for $\epsilon > 0$. Put

$$\psi_n(h) = \frac{1}{h^{1+\alpha p'-\epsilon}} \sum_{-n}^{n} | c_m |^{p'} | \sin mh |^{p'};$$

then the inequality

$$\Phi(h) = \left(\frac{1}{2\pi}\int_0^{2\pi} | \Delta f |^p d\theta\right)^{p'/p} \geq 2^{p'} \sum_{-\infty}^{\infty} | c_m \sin mh |^{p'}$$

implies the existence of the limit

$$\lim_{n\to\infty} \int_0^{2\pi} \psi_n(h)\, dh = \lim_{n\to\infty} \sum_{-n}^{n} | c_m |^{p'} \int_0^{2\pi} \frac{| \sin mh |^{p'}}{h^{1+\alpha p'-\epsilon}} dh,$$

since (2.3) is finite. There is a positive constant A satisfying

$$\int_0^{2\pi} \frac{|\sin mh|^{p'}}{h^{1+\alpha p'-\epsilon}}\, dh > A\,|m|^{\alpha p'-\epsilon} \qquad (|m| \geq 1).$$

Hence the series

$$\sum_{-\infty}^{\infty} |c_m|^{p'}\,|m|^{\alpha p'-\epsilon} \tag{2.4}$$

is convergent. This completes the proof of the lemma.

To demonstrate the theorem of Titchmarsh, we may obviously suppose that

$$\frac{p}{p+\alpha p-1} = \frac{p'}{1+\alpha p'} < k < \frac{p}{p-1};$$

then letting $(\alpha k-\epsilon)p'/(p'-k) > 1$, $\epsilon > 0$, $s = \alpha k - \epsilon$, we have

$$\begin{aligned} \sum_{-\infty}^{\infty} |c_m|^k &= \sum_{-\infty}^{\infty} (|c_m|^k\,|m|^s)(|m|^{-s}) \\ &\leq \Big(\sum_{-\infty}^{\infty} |c_m|^{p'}\,|m|^{sp'/k}\Big)^{k/p'}\Big(\sum_{-\infty}^{\infty} |m|^{-sp'/(p'-k)}\Big)^{(p'-k)/p'}. \end{aligned} \tag{2.5}$$

The last series of (2.5) is convergent, since

$$\frac{sp'}{p'-k} > 1.$$

Further, observing

$$sp'k^{-1} = (\alpha k-\epsilon)p'k^{-1} = \alpha p' - \epsilon' \qquad (\epsilon' = p'k^{-1}\epsilon > 0),$$

the convergence of $\sum |c_m|^k$ follows from the second part of Lemma 1.

3. **Proof of Theorem** 1. We require the following lemmas.

Lemma 2. *If $f(x) \in \mathrm{Lip}\,(\alpha, p)$, $0 < \alpha \leq 1$, $q > p > 1$, $\alpha - 1/p + 1/q > 0$, then*

$$f(x) \in \mathrm{Lip}\,(\alpha - p^{-1} + q^{-1}, q).$$

Lemma 3. *If $f \in \mathrm{Lip}\,(\alpha, p)$, then $f \in L^p$.*

Lemma 4. *If $f \in \mathrm{Lip}\,(\alpha, p)$, $p \geq 1$, $1 \geq \alpha > 0$, $f(\theta) \sim \sum c_m e^{im\theta}$, then $c_m = O(|m|^{-\alpha})$.*

These three lemmas are given in [5].

The existence of the integral

$$\int_{-\pi}^{\pi} f_1(\xi)f_2(\xi+x)\,d\xi$$

can be seen by Hölder's inequality. In fact, from Lemma 2, we have

$$f_1 \varepsilon \text{ Lip } (\alpha - p^{-1} + r^{-1}, r), \qquad f_2 \varepsilon \text{ Lip } (2^{-1}p^{-1} - q^{-1} + s^{-1}, s),$$

where $p < r \leq 2p$ and $s \leq 2p/(2p - 1)$. By Lemma 3 we have

$$f_1 \varepsilon L^{2p}, \qquad f_2 \varepsilon L^{2p/(2p-1)}.$$

Let

$$f_1 \sim \sum c_m e^{im\theta}, \qquad f_2 \sim \sum d_m e^{im\theta}.$$

Then

$$f(x) \sim \sum c_m d_m e^{im\theta}.$$

Then

$$\sum_{-N}^{N} | c_m d_m | \leq \{\sum_{-N}^{N} | c_m m^{\alpha'} |^{p'}\}^{1/p'} \{\sum_{-N}^{N} | d_m m^{-\alpha'} |^{p}\}^{1/p}.$$

If we choose α' such that $\alpha > \alpha' > 1/2p$, then the right side of the above inequality is bounded, as $N \to \infty$, by Lemma 1 and Lemma 4. Theorem 1 is thus proved.

4. **Proof of Theorem** 4. To prove Theorem 4, we appeal to the function $\Phi(h)$ of (2.2). First, we prove two lemmas.

LEMMA 5. *The existence of the integral*

$$J = \int_0^{2\pi} \Phi(h)(h^{p'})^{-1} | \log h^{-1} |^{(\alpha+R)p'-\epsilon} dh \tag{4.1}$$

$$(\epsilon > 0, R = p^{-1} - (p')^{-1} \geq 0)$$

implies the convergence of the series

$$\sum_{-\infty}^{\infty}{}' | c_m |^{p'} | m |^{p'-1}(\log | m |)^{(\alpha+R)p'-\epsilon}. \tag{4.2}$$

In fact, from Hausdorff's inequality, the existence of the integral (4.1) implies that of

$$\lim_{n\to\infty} \sum_{-n}^{n} | c_m |^{p'} \int_0^{2\pi} | \sin mh |^{p'}(h^{p'})^{-1} | \log h^{-1} |^{(R+\alpha)p'-\epsilon} dh. \tag{4.3}$$

Now the integral in (4.3) for $| m | > 0$ is greater than

$$\int_0^{\pi/4m} | \sin mh |^{p'}(h^{p})^{-1}(\log h^{-1})^{(R+\alpha)p'-\epsilon} dh$$

$$> (2\pi^{-1})^{p'} | m |^{p'} \int_0^{\pi/4|m|} (\log h^{-1})^{(R+\alpha)p'-\epsilon} dh$$

$$> A | m |^{p'-1}(\log | m |)^{(R+\alpha)p'-\epsilon}.$$

Therefore, (4.2) is convergent, since (4.3) exists.

LEMMA 6. *If $f(\theta)$ satisfies* (1.7), *then the integral* (4.1) *exists.*

Let $0 < \delta < 2\pi$, and write

$$J = \int_0^{\delta} \Phi(h)(h^{p'})^{-1}(\log h^{-1})^{(R+\alpha)p'-\epsilon} + \int_{\delta}^{2\pi} \Phi(h)(h^{p'})^{-1} \mid \log h^{-1} \mid^{(\alpha+R)p'-\epsilon}$$

$$= J_1 + J_2 .$$

Evidently, $J_2 = O(1)$. And

$$J_1 = O\left(\int_0^{\delta} h^{p'/p-p'}(\log h^{-1})^{(R+\alpha)p'-\epsilon-p'/p-\alpha p'}\, dh\right)$$

$$= O\left(\int_0^{\delta} h^{-1}(\log h^{-1})^{-1-\epsilon}\, dh\right) = O(1)$$

since $\Phi(h) = O(h^{p'/p}(\log h^{-1})^{-p'/p-\alpha p'})$, by (1.7).

To prove the first part of Theorem 4, we write

$$\sum{}' \mid c_m \mid (\log \mid m \mid)^T = \sum{}' \mid c_m \mid \mid m \mid^{(p'-1)/p'}(\log \mid m \mid)^{S+T} \mid m \mid^{-1/p}(\log \mid m \mid)^{-S}.$$

By Hölder's inequality, the right member is not greater than

$$[\sum \mid c_m \mid^{p'} \mid m \mid^{p'-1}(\log \mid m \mid^{(S+T)p'})]^{1/p'}[\sum \mid m \mid^{-1}(\log \mid m \mid)^{-Sp}]^{1/p}.$$

If $T < R + \alpha - 1/p$ or $R + \alpha - T > 1/p$, then we can choose a positive number ϵ such that $S = R + \alpha - T - \epsilon > 1/p$, so that

$$(S + T)p' = (R + \alpha - \epsilon)p' < (R + \alpha)p'.$$

Therefore, $\sum' \mid c_m \mid (\log \mid m \mid)^T$ converges for $T < R + \alpha - 1/p$, by Lemma 5 and Lemma 6.

The following example suffices to prove the second part of Theorem 4. Let $1 < p \leq 2$, $b > 0$, $ap > 1$, $(b - a)p > -1$, and

$$f(\theta) = (\log \mid \theta \mid^{-1})^{-a}(1 - \mid \theta \mid)^b.$$

Then we have

$$\int_{-\pi}^{\pi} \mid \Delta f \mid^p d\theta = \int_{-\pi}^{\pi} \mid f(\theta + h) - f(\theta - h) \mid^p d\theta = O(h(\log h^{-1})^{-ap}).$$

Put $ap = 1 + \alpha p$; then $\alpha = a - 1/p > 0$, since $ap > 1$. Let c_m be the Fourier coefficient of f. We have

$$2\pi c_m = \int_{-\pi}^{\pi} (1 - |\theta|)^b (\log |\theta|^{-1})^{-a} e^{-im\theta}\, d\theta$$

$$= m^{-1} \int_{-m\pi}^{m\pi} (1 - |\theta m^{-1}|)^b (\log |\theta m^{-1}|)^{-a} e^{i\theta}\, d\theta$$

$$= m^{-1}(\log |m|)^{-a} \int_{-m\pi}^{m\pi} (1 - |\theta m^{-1}|)^b (1 - \log |\theta|/\log |m|)^{-a} e^{-i\theta}\, d\theta.$$

Hence there is a positive constant A such that

$$|c_m| > A\, |m|^{-1}(\log |m|)^{-a},$$

so that

$$|c_m|\, (\log |m|)^{R+\alpha-1/p} > A\, |m|^{-1}(\log |m|)^{R+\alpha-1/p-a} = A(|m| \log |m|)^{-1}.$$

This completes the proof of Theorem 4.

If we put $a = 1$ in the foregoing example, the second part of Theorem 2 is proved. Theorem 3 can be proved with similar arguments.

Bibliography

1. S. Bernstein, *Sur la convergence absolue des séries trigonométriques*, Comptes Rendus, vol. 158(1914), pp. 1661–1663.
2. S. Bernstein, *Sur la convergence absolue des séries trigonométriques*, Comptes Rendus, vol. 199(1934), pp. 397–400.
3. K. K. Chen, *On the class of functions with absolutely convergent Fourier series*, Proceedings of the Imperial Academy, vol. 4(1928), p. 517.
4. G. H. Hardy and J. E. Littlewood, *Notes on the theory of series* (IX): *on the absolute convergence of Fourier series*, Journal of the London Mathematical Society, vol. 3 (1928), pp. 250–253.
5. G. H. Hardy and J. E. Littlewood, *A convergence criterion for Fourier series*, Mathematische Zeitschrift, vol. 28(1928), pp. 612–634.
6. R. Salem, *Sur certaines fonctions continues et les propriétés de leurs séries de Fourier*, Comptes Rendus, vol. 201(1935), pp. 703–705.
7. O. Szász, *Ueber den Konvergenzexponenten der Fourierschen Reihen gewisser Funktionenklassen*, Münchner Sitzungsberichte, 1922, pp. 135–150.
8. E. C. Titchmarsh, *A note on Fourier transforms*, Journal of the London Mathematical Society, vol. 2(1927), pp. 148–150.
9. W. H. Young, *On a class of parametric integrals and their application in the theory of Fourier series*, Proceedings of the Royal Society, (A), vol. 25(1911), p. 401.
10. A. Zygmund, *Remarque sur la convergence absolue des séries de Fourier*, Journal of the London Mathematical Society, vol. 3(1928), pp. 194–196.

THE CALCULUS OF VARIATIONS IN ABSTRACT SPACES

By H. H. Goldstine

The classical non-parametric problem of the calculus of variations deals with arcs defined by functions $y_p(x)$ $(p = 1, 2, \cdots, n)$. In the present paper we allow the parameter p to range over a quite arbitrary set $\mathfrak{P}$, and seek conditions that an arc $y_p(x)$ render an integral of the form

$$I = \int_{x_1}^{x_2} f[x, y(x), y'(x)]\, dx$$

a minimum in a class of admissible arcs. It is shown that this more general problem has a theory as complete as that of the classical problem. In a subsequent paper the author will take up the problem of Bolza in this general environment. In the first six sections, analogues of the familiar four necessary conditions are obtained. In §7 the sufficiency proofs are made, and in §8 the relation beteen conjugate points and the positiveness of the second variation is discussed.

1. **Formulation of the problem.** We shall use the notation $\mathfrak{R}$ to represent the set of real numbers, $\mathfrak{P}$ an arbitrary class of elements p, and $\mathfrak{B}$, an arbitrary Banach space of functions v on $\mathfrak{P}$ to $\mathfrak{R}$. It will be supposed that $(\mathfrak{R}, \mathfrak{B}, \mathfrak{B})_0$ is a region of the composite space $(\mathfrak{R}, \mathfrak{B}, \mathfrak{B})$ of sets (r, v, w) and that f on $(\mathfrak{R}, \mathfrak{B}, \mathfrak{B})_0$ to $\mathfrak{R}$ is a function of class C^{iv} uniformly on $(\mathfrak{R}, \mathfrak{B}, \mathfrak{B})_0$. See [4], [5], [8]. An *admissible arc* $y(x)$ is a continuous function y on (x_1, x_2) to $\mathfrak{B}$ which consists of a finite number of pieces on each of which $y'(x) \equiv \delta_x y(x; 1)$ exists and is continuous [5; 164] and such that each set $(x, y(x), y'(x))$ is in the fundamental region $(\mathfrak{R}, \mathfrak{B}, \mathfrak{B})_0$. An *admissible variation* is a function on an interval (x_1, x_2) having the continuity and differentiability properties of an admissible arc.

Our problem may then be formulated as that of finding in the class of admissible arcs joining two fixed points (x_1, y_1) and (x_2, y_2) one which minimizes the integral

$$I(C) = \int_{x_1}^{x_2} f[x, y(x), y'(x)]\, dx. \tag{1.1}$$

To carry through our analysis we shall suppose that there exists a mapping $(v_p \mid p \in \mathfrak{P})$ of $\mathfrak{P}$ onto a bounded subset $\mathfrak{B}_1$ of $\mathfrak{B}$ such that the linear extension of $\mathfrak{B}_1$ is dense in $\mathfrak{B}$, the limits being taken in the Moore-Smith sense; i.e., to each v in $\mathfrak{B}$ there corresponds a set of real numbers $a_{p\tau}$, where τ is a finite subset of

Received April 20, 1942. Presented to the American Mathematical Society, April 11, 1941, under the title *The parametric problems of the calculus of variations in general analysis.*

$\mathfrak{P}$, such that $v = \lim_{\tau} \sum_{q \in \tau} a_\tau \nu_q$, it being understood that $a_\tau = (a_{p\tau} \mid p \in \mathfrak{P})$. It is also assumed that $f(x, y + a\nu_p, y' + b\nu_p)$ is in $\mathfrak{B}$ for a, b sufficiently small.

Calculating the first and second differentials of the functional (1.1) along the minimizing arc E we find that the relations

$$
\begin{aligned}
I_1(\eta) = \int_{x_1}^{x_2} \{ &\delta_y f[x, y(x), y'(x); \eta(x)] \\
&+ \delta_{y'} f[x, y(x), y'(x); \eta'(x)]\}\, dx = 0,
\end{aligned} \tag{1.2}
$$

$$
\begin{aligned}
I_2(\eta) = \int_{x_1}^{x_2} \{ &\delta^2_{yy} f[x, y(x), y'(x); \eta(x), \eta(x)] \\
&+ 2\delta^2_{yy'} f[x, y(x), y'(x); \eta(x), \eta'(x)] \\
&\quad + \delta^2_{y'y'} f[x, y(x), y'(x); \eta'(x), \eta'(x)]\}\, dx \geq 0
\end{aligned} \tag{1.3}
$$

hold for all admissible variations η which are such that $\eta(x_1) = \eta_2(x) = 0_v$, the zero element in $\mathfrak{B}$.

2. **First necessary conditions.** If in the equation (1.2) one replaces $\eta(x)$ by the special variation $\nu_p\zeta(x)$, where ν_p is an arbitrary member of the family described directly below (1.1) and $\zeta(x)$ is a real-valued function vanishing at x_1 and x_2 , the well-known fundamental lemma of the calculus of variations shows that there must exist a $c = (c_p)$ on $\mathfrak{P}$ to $\mathfrak{R}$ such that the equation

$$
\delta_{y'} f(x; \nu_p) = \int_{x_1}^{x} \delta_y f(s; \nu_p)\, ds + c_p \tag{2.1}
$$

holds at each point of a minimizing arc, where the symbol $\delta_y f(x; \nu_p)$ is an abbreviation for $\delta_y f(x, y(x), y'(x); \nu_p)$ and $\delta_{y'} f(x; \nu_p)$ has a similar meaning. It is evident that $c_p = \delta_{y'} f(x_1 ; \nu_p)$ and hence that c_p is in $\mathfrak{B}$.

THEOREM 2.1. *A necessary and sufficient condition that the first variation* $I_1(\eta)$ *vanish for all admissible variations which vanish at* x_1 *and* x_2 *is that there exist a* c *in* $\mathfrak{B}$ *such that equation* (2.1) *holds.*

From what has been shown above it is necessary only to prove that this condition is sufficient for $I_1(\eta) = 0$. Let us observe first that there is a set of functions $a_\tau(x) = a_{p\tau}(x)$, where p ranges over $\mathfrak{P}$ and τ ranges over all finite subsets of $\mathfrak{P}$, such that

$$
\eta'(x) = \lim_{\tau} \sum_{p \in \tau} a_\tau \nu_p
$$

because of the density property of the set $[\nu_p]$. If we multiply both sides of (2.1) by $a_\tau(x)$, sum and pass to the limit, we find that the equation

$$
\delta_{y'} f(x; \eta'(x)) = \int_{x_1}^{x} \delta_y f(s; \eta'(x))\, ds + \delta_{y'} f(x_1 ; \eta'(x)) \tag{2.2}
$$

must hold. An analogue of the usual integration by parts shows that

$$\int_{x_1}^{x_2} dx \int_{x_1}^{x} \delta_y f(s; \eta'(x))\, ds = \int_{x_1}^{x_2} \delta_y f(s; \eta(x_2))\, ds - \int_{x_1}^{x_2} \delta_y f(x; \eta(x))\, dx,$$

as may be seen by interchanging the order of integration in the left side. It follows therefore that

$$0 = I_1(\eta) - \int_{x_1}^{x_2} \delta_{y'} f(x_1\, ; \eta'(x))\, dx = I_1(\eta)$$

for all η such that $\eta(x_1) = \eta(x_2) = 0$, , since

$$\int_{x_1}^{x_2} \delta_{y'} f(x_1\, ; \eta'(x))\, dx = \delta_{y'} f(x_1\, ; \eta(x))\Big|_{x_1}^{x_2}.$$

Corollary 2.1. *If $I_1(\eta)$ is computed along an arc without corners, then $I_1(\eta)$ vanishes for all admissible variations such that $\eta(x_1) = \eta(x_2) = 0$, if and only if the equation*

$$\frac{d}{dx}\, \delta_{y'} f(x; \nu_p) = \delta_y f(x; \nu_p) \tag{2.3}$$

holds.

If the arc along which I_1 is computed has no corners, then clearly equations (2.1) and (2.3) are equivalent.

It is of some interest to calculate $I_1(\eta)$ along an arc E satisfying (2.3) for a variation η that does not necessarily vanish at x_1 and x_2. To do this we note with the help of formula (2.2) and those following it that

$$I_1(\eta) = \delta_{y'} f(x_1\, ; \eta(x))\Big|_{x_1}^{x_2} + \int_{x_1}^{x_2} \delta_y f(x; \eta(x_2))\, dx.$$

It follows at once from equation (2.3) and the property of the set ν_p that d/dx $\delta_{y'} f(x; \eta(x_2)) = \delta_y f(x; \eta(x_2))$, and hence that

$$I_1(\eta) = \delta_{y'} f(x; \eta(x))\Big|_{x_1}^{x_2}. \tag{2.4}$$

Corollary 2.2. *At each point of a minimizing arc the corner condition*

$$\lim_{\xi < x} \delta_{y'} f[x, y(x), y'(\xi); \nu_p] = \lim_{x < \eta} \delta_{y'} f[x, y(x), y'(\eta); \nu_p]$$

must hold for each p in $\mathfrak{P}$.

With the aid of the differentiability hypothesis we see that for each ζ in $\mathfrak{B}$ the function $\delta^2_{y'y'} f(x, y, y'; \nu_p, \zeta)$ is in $\mathfrak{B}$; and hence that $(\delta^2_{y'y'} f(x, y, y'; \nu_p, \zeta) \mid p \,\varepsilon\, \mathfrak{P})$ is, for fixed x, y, y', a mapping of $\mathfrak{B}$ onto itself. If at a point (x, y, y') this mapping has a reciprocal [8; 145], the point (x, y, y') is said to be *non-singular*. An

arc E is *non-singular* if it is non-singular at each of its points. Using this definition we can state an analogue of Hilbert's differentiability condition.

COROLLARY 2.3. *Near each non-singular point of a minimizing arc E, which is not a corner, the function $y(x)$ defining E is of class C^{iv}.*

To prove this condition we observe that the equation

$$\delta_{v} \cdot f[x, y(x), v, \nu_p] - \int_{x_1}^{x} \delta_y f[s, y(s), y'(s); \nu_p]\, ds - c_p = 0,$$

has an initial solution $(x, v) = (x, y'(x))$ at which the differential with respect to v has a reciprocal. It follows at once from known implicit-function theorems [8; 150] that the unique solution $v(x) = y'(x)$ of this equation is of class C'''.

3. **The extremals.** An arc which is of class C'' and satisfies the equation (2.3) is called an extremal and evidently satisfies the equation

$$\begin{aligned} \delta^2_{xy'} \cdot f(x, y, y'; 1, \nu_p) + \delta^2_{yy'} \cdot f(x, y, y'; y', \nu_p) + \delta^2_{y'y'} \cdot f(x, y, y'; y'', \nu_p) \\ - \delta_y f(x, y, y'; \nu_p) = 0_\cdot\,. \end{aligned} \tag{3.1}$$

It is clear that every non-singular sub-arc without corners in a minimizing arc is an extremal. We have then the following imbedding theorem.

THEOREM 3.1. *Every non-singular extremal arc E is imbedded for values $x_1 \leq x \leq x_2$, $(a, b) = (a_0, b_0)$ in a family of extremals $y = y(x, a, b)$ with (a, b) in $(\mathfrak{B}, \mathfrak{B})$ such that y, y_x are of class C'' uniformly on a neighborhood of the values (x, a_0, b_0) belonging to E. At some point x_0 the identities*

$$y(x_0, a, b) = a, \qquad y_x(x_0, a, b) = b \tag{3.2}$$

are valid.

To prove the theorem let $\mathfrak{Y}$ be the class of all arcs continuous on (x_1, x_2), let $\mathfrak{Z}$ be the Cartesian product $(\mathfrak{Y}, \mathfrak{Y})$, and let $\mathfrak{X} = (\mathfrak{B}, \mathfrak{B})$. The spaces $\mathfrak{Y}$ and $\mathfrak{Z}$ are both Banach spaces of the type A_0 of Graves, and we may use one of his imbedding theorems [4; 540]. Let $F(y'', z)$ with $z = (y, y')$ denote the left member of (3.1) and let $H(x, y'')$ denote the pair $z = (a + b(r - r_0) + \int_{r_0}^{r}\int_{r_0}^{t} y''(s)ds\, dt$, $b + \int_{r_0}^{r} y''(s)ds)$, where $x = (a, b)$ and r_0 is a fixed constant. It is not difficult to show that the hypotheses of Graves' theorem are satisfied and hence that there is a solution $y''(r, a, b)$ of the equation $F[y'', H(a, b, y'')] = 0_\cdot$, which is defined on a neighborhood of $(a_0, b_0) = [y(r_0), y'(r_0)]$ and with values in a

neighborhood of those belonging to $y''(r)$. This solution is of class C'' uniformly on its domain and hence

$$y(x, a, b) \equiv a + b(x - x_0) + \int_{x_0}^{x} \int_{x_0}^{t} y''(s, a, b)\, ds\, dt$$

is effective in the theorem.

It will be shown later that the linear operator

$$(3.3) \qquad \Delta(x, \alpha, \beta) \equiv (\delta y(x, a_0, b_0; \alpha, \beta),\ \delta y_x(x, a_0, b_0; \alpha, \beta)),$$

where δ indicates differentiation with respect to (a, b), has a reciprocal at each value x on (x_1, x_2).

4. **The Jacobi condition.** Hereafter we shall designate by $2\omega(x, \eta, \eta')$ the integrand of the integral (1.3). As usual we consider the problem of minimizing the functional I_2 of equation (1.3) in the class of all admissible variations vanishing at x_1 and x_2, and we remark that $\eta \equiv 0_v$ is a minimizing arc. It will be assumed that I_2 has been computed along a non-singular extremal without corners. The extremals for our new problem must then satisfy the equation

$$(4.1) \qquad \frac{d}{dx} \delta_{\eta'}\omega(x; \eta(x), \eta'(x); \nu_p) = \delta_\eta \omega(x, \eta(x), \eta'(x); \nu_p)$$

and be non-singular. Known existence theorems [4; 547] tell us that this equation has a unique solution through each set (x_0, η_0, η_0'). An extremal for the *accessory* problem formulated above must also satisfy the corner condition of Corollary 2.2, which implies that

$$\lim_{\xi < x < \zeta} \delta^2_{y'y'} f(x; \eta'(\xi) - \eta'(\zeta), \nu_p) = 0_v\,.$$

Since, however, the original arc is non-singular, it follows that every accessory extremal is without corners.

A value x_6 is said to be *conjugate* to x_1 if there is an accessory extremal, i.e., an extremal for the accessory problem, vanishing at x_1 and x_6 but not identically zero on (x_1, x_6). We then have

THEOREM 4.1. *For a non-singular minimizing arc E_{12} without corners there can be no value x_6 between x_1 and x_2 conjugate to x_1.*

Let $\eta(x)$ be an accessory extremal defining a conjugate value x_6 and let $u(x) \equiv \eta(x)$ on (x_1, x_6) and $\equiv 0_v$ on (x_6, x_2). This arc minimizes the functional $I_2(\eta)$; hence it satisfies the corner condition at x_6, whence it follows that $\eta(x_6) = \eta'(x_6) = 0_v$. It then follows that $\eta(x) \equiv 0_v$.

THEOREM 4.2. *For a family of extremal arcs $y = y(x, a, b)$ of the sort described in Theorem 3.1 the linear and continuous operator Δ in (3.3) has a reciprocal at every value of x on the interval (x_1, x_2).*

To prove this result, we substitute the given family into the equation and differentiate with respect to (a, b). It follows at once that $\delta y(x, a_0, b_0; \alpha, \beta)$ is a solution of the accessory equation (4.1), and with the help of equations (3.2) that $\Delta(x_0, \alpha, \beta)$ has a reciprocal. Consider an arbitrary pair v, w of elements in $\mathfrak{B}$ and an arbitrary value x_6 of x. It is known that the accessory equation (4.1) has a unique solution $\eta(x)$ such that $\eta(x_6) = v$, $\eta'(x_6) = w$. The equations

$$\delta y(x_0, a_0, b_0; \alpha, \beta) = \eta(x_0), \qquad \delta y_x(x_0, a_0, b_0; \alpha, \beta) = \eta'(x_0)$$

have a solution $(\alpha, \beta) = (\alpha_0, \beta_0)$ and hence the function $\eta(x) - \delta y(x, a_0, b_0; \alpha_0, \beta_0)$ is a solution of the accessory equation, which vanishes at x_0 along with its derivative. Accordingly, $\eta(x) \equiv \delta y(x, a_0, b_0; \alpha_0, \beta_0)$, and it follows that at the arbitrary value x_6, $\Delta(x, \alpha, \beta)$ assumes at least once every value in the composite space $(\mathfrak{B}, \mathfrak{B})$. If for some value x_6 the equation $\Delta(x_6, \alpha, \beta) = 0_{vv}$ had a solution $(\alpha, \beta) \neq (0_v, 0_v)$, then an argument like the one above shows that $\Delta(x_0, \alpha, \beta) = 0_{vv}$, which is a contradiction. It has now been shown that $\Delta(x, \alpha, \beta)$ has an inverse on $(\mathfrak{B}, \mathfrak{B})$ to $(\mathfrak{B}, \mathfrak{B})$ and it remains only to show that it is bounded. This follows at once from a well-known result [1].

Corollary 4.1. *For every solution $\eta(x)$ of the accessory equation there exists an element (α, β) in $(\mathfrak{B}, \mathfrak{B})$ such that $\eta(x) \equiv \delta y(x, a_0, b_0; \alpha, \beta)$.*

To develop a theory of fields, we find it desirable to introduce the notion of a *complete arc*. An arc E_{12} is complete if every pair of points (x_1, y_1), (x_6, y_6) in xy-space, with x_6 not conjugate to x_1, can be joined by at least one accessory extremal. The purpose of this condition is to insure that the linear operator

$$D(x_1, x, \alpha, \beta) \equiv (\delta y(x_1, a_0, b_0; \alpha, \beta), \delta y(x, a_0, b_0; \alpha, \beta)) \tag{4.2}$$

takes on, at least once, every value (ρ, σ) in $(\mathfrak{B}, \mathfrak{B})$ when x is not conjugate to x_1. It is assumed throughout the next two sections that E_{12} is a complete, nonsingular extremal arc.

Theorem 4.3. *The values conjugate to x_1 are determined by the values of x for which the linear, continuous operator $D(x_1, x, \alpha, \beta)$ defined above in* (4.2) *fails to have a reciprocal.*

To establish the theorem we find it convenient first to show that the conjugate values are determined by the values of x for which D has no inverse; i.e., the values of x for which the equation $D(x_1, x, \alpha, \beta) = (0_v, 0_v)$ has a solution $(\alpha, \beta) \neq (0_v, 0_v)$.

If $\eta(x)$ is the accessory extremal defining a value x_6 conjugate to x_1, we can determine values $(\alpha, \beta) \neq (0_v, 0_v)$ so that $\eta(x) \equiv \delta y(x, a_0, b_0; \alpha, \beta)$. Since $\eta(x_1) = \eta(x_6) = 0_v$, it is clear that $D(x_1, x_6, \alpha, \beta)$ cannot have an inverse. Conversely, if $D(x_1, x_6, \alpha, \beta)$ does not have an inverse, constants $(\alpha, \beta) \neq (0_v, 0_v)$ can be so determined that $\eta(x) \equiv \delta y(x, a_0, b_0; \alpha, \beta)$ vanishes at x_1 and x_6, but is not identically zero on (x_1, x_6). If it were identically zero, then

$\Delta(x, \alpha, \beta)$ in (3.3) could not have a reciprocal, which is a contradiction of Theorem 4.2.

It remains only to show that D has an inverse if and only if it has a reciprocal. If D has an inverse for a value x, then x is not conjugate to x_1, as we have just seen. With the help of a theorem of Banach [1], it follows that D has a reciprocal. The converse is obvious.

LEMMA 4.1 *In a sufficiently small neighborhood of x_1 there is no value x conjugate to x_1.*

The operator $D(x_1, x, \alpha, \beta)$ has a reciprocal for $x \neq x_1$ if and only if the operator

$$A(x_1, x, \alpha, \beta) \equiv \{\delta y(x_1, a_0, b_0; \alpha, \beta), [\delta y(x, a_0, b_0; \alpha, \beta) - \delta y(x_1, a_0, b_0; \alpha, \beta)]/(x - x_1)\} \tag{4.3}$$

has one. Using Taylor's theorem [5; 173] we may express A in the form

$$(\delta y(x_1, a_0, b_0; \alpha, \beta), \int_0^1 \delta y_x(x_1 + r(x - x_1), a_0, b_0; \alpha, \beta)\, dr), \tag{4.4}$$

and hence $A(x_1, x_1, \alpha, \beta) = \Delta(x_1, x_1, \alpha, \beta)$ has a reciprocal. It follows easily from this fact that $A(x_1, x, \alpha, \beta)$ has a reciprocal for x near x_1.

THEOREM 4.4. *If there is no value conjugate to x_1 on the interval $x_1 < x \leq x_2$, then there is an $x_0 < x_1$ such that there is no value conjugate to x_0 on (x_1, x_2).*

The operator $A(\xi, x, \alpha, \beta)$ defined in (4.3) has a reciprocal for $\xi = x_1$, $x_1 \leq x \leq x_2$, as shown above. It will therefore have a reciprocal in a neighborhood of each value (x_1, x), and by the Borel theorem there is then a uniform neighborhood of the sets (x_1, x) with $x_1 \leq x \leq x_2$ in which A has a reciprocal. The theorem follows from this fact since $D(x_0, x, \alpha, \beta)$ has a reciprocal for $x_1 \leq x \leq x_2$ if and only if $A(x_0, x, \alpha, \beta)$ does.

THEOREM 4.5. *Let $y = y(x, a)$ $(a \in \mathfrak{B})$ be a family of extremals of class C'' uniformly on its domain, containing E_{12} for $a = a_0$ and such that all the extremals pass through the point 1. Then the values conjugate to x_1 are determined by the values of x for which the function $\delta y(x, a_0; \alpha)$ fails to have a reciprocal, provided $\delta y_x(x_1, a_0; \alpha)$ has a reciprocal.*

We first show that every accessory extremal $\eta(x)$ vanishing at x_1 is expressible as $\delta y(x, a_0; \alpha)$. We can determine an α in $\mathfrak{B}$ such that $\delta y(x_1, a_0; \alpha) = 0_v$, $\delta y_x(x_1, a_0; \alpha) = \eta'(x)$. The arc $\eta(x) - \delta y(x, a_0; \alpha)$ is then identically zero. The remainder of the proof is quite like that of Theorem 4.3.

5. **Simply covered regions.** We consider in this section a family

$$y = y(x, a) \qquad (x_1 - e < x < x_2 + e, \|a\| < e) \tag{5.1}$$

of extremals containing E_{12} for $x_1 \leq x \leq x_2$, $a = a_0$, and defined by a function of class C'' uniformly on its domain. This family simply covers a region $\mathfrak{F}$ of xy-space if there is a unique extremal through each point (x, y) of $\mathfrak{F}$. If $\mathfrak{F}$ is simply covered, we can solve equation (5.1) for a and get a single-valued solution $A(x, y)$. The function

$$p(x, y) \equiv y_x(x, A(x, y)) \tag{5.2}$$

is called the slope-function of the family.

LEMMA 5.1. *If the linear operator $\delta y(x, a_0 ; \alpha)$ has a reciprocal at each value x on (x_1, x_2), then there is a region $\mathfrak{F}$ of xy-space that is simply covered by the family* (5.1) *and that contains E_{12}. In this region the slope-function p is of class C'.*

To prove the lemma we apply an extended implicit-function theorem of Graves [4; 532] to the equation

$$y - y(x, a) = 0_x . \tag{5.3}$$

This equation has as an initial solution the set $\mathfrak{Z}^0$ of all (x, y, a) such that $x_1 \leq x \leq x_2$, $y = y(x, a)$, $a = a_0$, which is self-compact. If we designate the xy-projection of this initial solution by $\mathfrak{U}^0$, then there are positive constants c, d and a solution A on $(\mathfrak{U}^0)_d$ to $\mathfrak{B}$—the notation $(\mathfrak{U}^0)_d$ means the sum of all neighborhoods $(u)_d$ of radius d of points u in $\mathfrak{U}^0$—such that, for every (x, y) in $\mathfrak{B}^0$, $A(x, y)$ is the unique solution (x, y, a) of (5.3) for which (x, y, a) is in a uniform neighborhood $(\mathfrak{Z}^0)_c$ of $\mathfrak{Z}^0$, and it is of class C'' on $(\mathfrak{U}^0)_d$. We may now choose the region $\mathfrak{F}$ to be the neighborhood $(\mathfrak{U}^0)_d$.

THEOREM 5.1. *If a non-singular extremal arc E_{12} has no value conjugate to x_1 on $x_1 < x \leq x_2$, then there is a value $x_0 < x_1$ such that the extremals through $(x_0, y(x_0))$ form a family* (5.1) *which simply covers a neighborhood $\mathfrak{F}$ of E_{12} in xy-space.*

If in Theorem 3.1 we set $a = y(x_0)$, we get a new family $y(x, b)$ of extremals all of which pass through $(x_0, y(x_0))$ and for which $\delta y(x, b_0 ; \beta)$ has a reciprocal. The theorem then follows from the lemma above.

Before proceeding further it is desirable to establish the following lemma.

LEMMA 5.2. *If a neighborhood $\mathfrak{F}$ of the arc E_{12} is simply covered by a family of extremals all of which pass through some fixed point (x_0, y_0), then for an arbitrary arc C_{12} in $\mathfrak{F}$ joining the ends of E_{12} it is true that*

$$I(E_{12}) = \int_{x_1}^{x_2} [f(x, y, p(x, y)) + \delta_{y'} f(x, y, p(x, y); Y' - p(x, y))] \, dx,$$

where (y, Y') are the elements belonging to C_{12} at the value x.

To establish the lemma let $y = y(x, t)$ $(t_1 \leq t \leq t_2)$ be the extremal through (x_0, y_0) and the point $[x(t), y(t)]$ on C_{12}. Calculating the derivative of

$$I(t) = \int_{x_0}^{x(t)} f[x, y(x, t), y_x(x, t)]\, dx$$

we find, with the help of (2.4), that

$$\begin{aligned} I'(t) &= f[x(t), y(t), p(x(t), y(t))]x'(t) \\ &\qquad + \delta_{y'} f[x(t), y(t), p(x(t), y(t)); y'(t) - p(x(t), y(t))x'(t)]. \end{aligned} \tag{5.4}$$

If $x_1 = x(t_1)$, $x_2 = x(t_2)$, we see that $y(x, t_1)$ defines E_{01} and that $y(x, t_2)$ defines E_{02} and hence that $I(E_{12}) = I(t_2) - I(t_1)$. The lemma then follows from (5.4) by means of an integration.

6. **The conditions of Weierstrass and Legendre.** To establish an analogue of the Weierstrass condition we use the method developed by Graves [6] for proving this condition. Let us consider a value x^* inside the interval (x_1, x_2) and a value Y' such that $(x^*, y(x^*), Y')$ is admissible. We define a comparison arc $C_{b,e}$ by the relations

$$\varphi(x, b, e) = \begin{cases} y(x^*) + (x - x^*)Y' & (x^* < x \leq be + x^*), \\ y(x^* + b) + \dfrac{x - x^* - b}{(1 - e)b}\,[y(x^* + b) - y(x^*) - beY'] & \\ & (be + x^* < x < b + x^*), \\ y(x) & (\text{elsewhere}). \end{cases}$$

For sufficiently small positive values of b, e this arc is admissible and we have

$$\begin{aligned} 0 &\leq \frac{1}{b}\,[I(C_{b,e}) - I(E)] \\ &= \frac{1}{b}\int_{x^*}^{x^*+be} [f(x, \varphi(x, b, e), \varphi_x(x, b, e)) - f(x, y(x), y'(x))]\, dx \\ &\qquad + \frac{1}{b}\int_{x^*+be}^{x^*+b} [f(x, \varphi(x, b, e), \varphi_x(x, b, e)) - f(x, y(x), y'(x))]\, dx, \end{aligned}$$

if it is assumed that E is a minimizing arc. Using the law of the mean for integrals and letting b approach zero, we find that

$$\begin{aligned} 0 &\leq e[f(x^*, y(x^*), Y') - f(x^*, y(x^*), y'(x^*))] \\ &\qquad + (1 - e)\left\{f\left[x^*, y(x^*), \frac{1}{1 - e}\,(y'(x^*) - eY')\right] - f(x^*, y(x^*), y'(x^*))\right\}. \end{aligned}$$

If both members of this inequality are divided by e and if e is made to approach zero, we find that

$$E(x, y, y', Y') \equiv f(x, y, Y') - f(x, y, y') - \delta_{y'}f(x, y, y'; Y' - y') \geq 0. \tag{6.1}$$

THEOREM 6.1. *If E is a minimizing arc, then the condition $E(x, y, y', Y') \geq 0$ must hold for each (x, y, y') belonging to E and to each $Y' \neq y'$ such that (x, y, Y') is admissible.*

This condition was proved above for x interior to (x_1, x_2). At x_1 and x_2 the condition follows by simple continuity considerations.

THEOREM 6.2. *At each point (x, y, y') of a minimizing arc the condition*

$$\delta^2_{y'y'}f(x, y, y'; \eta, \eta) \geq 0 \tag{6.2}$$

must hold for every $\eta \in \mathfrak{B}$ such that $\| \eta \| = 1$.

To prove this we expand the E-function of (6.1) by means of Taylor's theorem and get

$$\begin{aligned} 0 &\leq E(x, y, y', Y') \\ &= \int_0^1 (1 - r)\delta^2_{y'y'}f[x, y, y' + (Y' - y')r; Y' - y', Y' - y'] \, dr. \end{aligned} \tag{6.3}$$

If we set $Y' = y' + e\eta$ in (6.3) and let e approach zero, we get condition (6.2).

7. **Sufficiency theorems.** As is customary we shall use the notations I, II, III, IV for the necessary conditions expressed in Theorems 2.1, 6.1, 6.2, 4.1. The condition II_N is the condition II extended to hold for all (x, y, y') in a neighborhood N of those belonging to E_{12}. The condition IV′ is the condition IV strengthened to exclude the possibility of x_2 being conjugate to x_1. Finally II'_N, III′ are the conditions II_N, III strengthened to exclude the equality signs.

LEMMA 7.1. *If a non-singular arc E satisfies condition II_N, then it also satisfies condition II'_N if N is small enough* [7].

If II'_N did not hold, there would be a set (x_0, y_0, y'_0, Y'_0) such that $y'_0 \neq Y'_0$ and $E(x_0, y_0, y'_0, Y'_0) = 0$, i.e., the function $g(y') = E(x_0, y_0, y', Y'_0)$ would attain its minimum at $y' = y'_0$. At this value then we would have [3]

$$0 = \delta g(y'_0; \eta) = -\delta^2_{y'y'}f(x_0, y_0, y'_0; \eta, Y'_0 - y'_0)$$

for all $\eta \in \mathfrak{B}$. This is impossible since the neighborhood N may be taken so small that in it $\delta^2_{y'y'}f(x_0, y_0, y'_0; \nu_p, \zeta)$ has a reciprocal.

THEOREM 7.1. *Let E_{12} be a complete, non-singular arc without corners satisfying conditions* I, III, IV′. *It then furnishes the integral I with a proper, weak relative minimum. If E_{12} satisfies the condition* II_N *instead of* III, *then the minimum is a strong relative one.*

It is clear that a non-singular arc satisfying III also satisfies III′ and the proof for a weak minimum is entirely analogous to the one for the classical case. The same remark applies to the proof of the strong minimum when one observes, with the help of Lemma 7.1, that II'_N is satisfied.

8. **The second variation.** It is assumed in this section that E_{12} is a complete, non-singular extremal arc. On the basis of what has preceded we shall develop an analogue of the Bliss-Clebsch [2] transformation of the second variation. To this end we study the accessory minimum problem, formulated in §4. The arc $\eta \equiv 0_*$ is evidently a complete, non-singular accessory extremal. If the arc E_{12} has no value conjugate to x_1 on the interval (x_1, x_2), then $\eta \equiv 0_*$ also has this property. If $y = y(x, b)$ is the family of extremals of Theorem 5.1, then the family $\eta = \delta y(x, b_0; \beta)$ simply covers the region of $x\eta$-space between the hyperplanes $x = x_1$, $x = x_2$ and all the arcs of the family pass through $(x_0, 0_*)$. The slope-functions of this family will be denoted by $\pi(x, \eta)$.

If $\eta(x)$ is an arbitrary admissible variation vanishing at x_1 and x_2, then Lemma 5.2 assures us that

$$I_2(\eta) = I_2(\eta) - I_2(0_*) = \int_{x_1}^{x_2} E_\omega[x, \eta(x), \pi(x, \eta(x)), \eta'(x)]\, dx, \tag{8.1}$$

where E_ω is the E-function computed for the integrand $2\omega(x, \eta, \eta')$ and has the value

$$E_\omega = \delta^2_{y'y'} f[x; \eta'(x) - \pi(x, \eta(x)), \eta'(x) - \pi(x, \eta(x))]. \tag{8.2}$$

THEOREM 8.1. *Let E_{12} be a complete, non-singular extremal arc which satisfies condition* III. *Then the second variation $I_2(\eta)$ is positive for all admissible variations $\eta(x) \not\equiv 0_*$ vanishing at x_1 and x_2 if and only if E_{12} satisfies condition* IV′.

If IV′ is satisfied, then $I_2(\eta) > 0$ follows from (8.1) and (8.2). To establish the converse, we see, with the help of Theorem 4.1, that $I_2(\eta) \geq 0$ implies that E_{12} satisfies condition IV. If x_2 is conjugate to x_1, then there is an accessory extremal $\eta(x)$ vanishing at x_1 and x_2 and not identically zero. By means of a formula analogous to (2.4) for 2ω, we see that $I_2(\eta) = 0$. Hence $I_2(\eta) > 0$ implies that IV′ is satisfied.

THEOREM 8.2. *Let E_{12} be a complete, non-singular extremal arc satisfying* III. *The second variation $I_2(\eta)$ is non-negative for all admissible variations $\eta(x)$ vanishing at x_1 and x_2 if and only if E_{12} satisfies condition* IV.

It suffices to show that IV is sufficient since the converse was proved above. Let $\eta_0(x)$ be an arbitrary admissible variation vanishing at x_1 and x_2 and let $H_n(x)$ be defined to be $\eta_0(x) - (x - x_1)\eta_0(x_2 - 1/n)/(x_2 - x_1 - 1/n)$ on $x_1 \leq x \leq x_2 - 1/n$ and 0_* on $x_2 - 1/n \leq x \leq x_2$. The variation H_n is admissible, vanishes at x_1 and x_2, and converges both in position and direction to η_0 on (x_1, x_2).

On the interval $(x_1, x_2 - 1/n)$ there is no value conjugate to x_1 and hence by Theorem 8.1 the integral

$$I_{2,n}(\eta) = \int_{x_1}^{x_2-1/n} 2\omega(x, \eta, \eta')\, dx$$

is positive for all $\eta \neq 0$, vanishing at x_1 and $x_2 - 1/n$. It is clear that $I_{2,n}(H_n)$ converges to $I_2(\eta_0)$ and that H_n vanishes at x_1 and $x_2 - 1/n$; hence $I_2(\eta_0) \geq 0$.

Bibliography

1. S. Banach, *Théorie des opérations linéaires*, p. 41.
2. G. A. Bliss, *The transformation of Clebsch in the calculus of variations*, Proceedings of the International Congress of Mathematicians, Toronto, 1924, vol. 1, p. 589.
3. H. H. Goldstine, *Minimum problems in the functional calculus*, Bulletin of the American Mathematical Society, vol. 46(1940), pp. 142–149.
4. Lawrence M. Graves, *Implicit functions and differential equations in general analysis*, Transactions of the American Mathematical Society, vol. 29(1927), pp. 514–552.
5. Lawrence M. Graves, *Riemann integration and Taylor's theorem in general analysis*, Transactions of the American Mathematical Society, vol. 29(1927), pp. 163–177.
6. Lawrence M. Graves, *The Weierstrass condition for multiple integral variation problems*, this Journal, vol. 5(1939), pp. 656–660.
7. M. R. Hestenes and W. T. Reid, *A note on the Weierstrass condition in the calculus of variations*, Bulletin of the American Mathematical Society, vol. 45(1939), pp. 471–473.
8. T. H. Hildebrandt and Lawrence M. Graves, *Implicit functions and their differentials in general analysis*, Transactions of the American Mathematical Society, vol. 29(1927), pp. 127–153.

University of Michigan.

THE POINT OF INFLEXION OF A PLANE CURVE

By Su-Cheng Chang

1. In attempting to generalize Bompiani's investigation for an inflexion of a plane curve [1] to the singularity of high order [3], the quartics of two cusps and an inflexion with concurrent cusp tangents and inflexion tangent are utilized here to represent several orders of the neighborhoods of the curve at its inflexion, as Bompiani has done by using cusped cubics. According as the inflexion is of the first or the second class, two types of canonical expansions are given, namely,

$$y = x^3 + \frac{27}{16}x^7 + qx^8 + (9),$$

$$y = x^3 + \frac{88}{3}x^8 + (9);$$

the former coincides with that given by Bompiani [1], while the latter is new. A similar improvement for the cusp of a plane curve has been established by the author [2].

2. There are quartics each of which has a point of inflexion at $O(0, 0, 1)$ and two cusps at $P_1(x_1, y_1, z_1)$ and $P_2(x_2, y_2, z_2)$. In the following lines we utilize a special kind of them whose cusp tangents and inflexion tangent are concurrent. If the inflexion tangent is taken for $y = 0$, then the equation of one of these quartics is

$$\begin{aligned}
&\left(\alpha\frac{x_1y - y_1x}{y_1} + \frac{\omega_1x + \omega_2y + \omega_3z}{y_1y_2}\right)^2 \\
&\times\left(\frac{\alpha^2}{\tau^2}y^2 - 2\frac{\alpha^2}{\tau}\frac{x_1y - y_1x}{y_1}y - 2\frac{\alpha}{\tau}y\cdot\frac{\omega_1x + \omega_2y + \omega_3z}{y_1y_2}\right) \\
&+ 4\alpha^3\left(\frac{x_1y - y_1x}{y_1}\right)^3\left(\frac{\omega_1x + \omega_2y + \omega_3z}{y_1y_2}\right) \\
&+ 6\frac{\alpha}{\tau}y\left(\alpha\frac{x_1y - y_1x}{y_1} + \frac{\omega_1x + \omega_2y + \omega_3z}{y_1y_2}\right)\left(\frac{\omega_1x + \omega_2y + \omega_3z}{y_1y_2}\right)^2 \\
&+ \alpha^4\left(\frac{x_1y - y_1x}{y_1}\right)^4 - 4\left(1 - \frac{\alpha^2\tau^3}{a}\right)\frac{\alpha}{\tau}y\left(\frac{\omega_1x + \omega_2y + \omega_3z}{y_1y_2}\right)^3 = 0,
\end{aligned} \tag{1}$$

Received April 22, 1942.

where we have placed

$$\omega_1 = y_1 z_2 - y_2 z_1, \qquad \omega_2 = z_1 x_2 - z_2 x_1, \qquad \omega_3 = x_1 y_2 - x_2 y_1,$$
$$\tau = \frac{y_1 y_2}{\omega_3} \qquad (y_1 y_2 \neq 0). \tag{2}$$

The cusp tangent at P_1 of the quartic is given by

$$\alpha\left(\frac{x_1 y - y_1 x}{y_1}\right) + \frac{\omega_1 x + \omega_2 y + \omega_3 z}{y_1 y_2} = 0. \tag{3}$$

Suppose that a curve C has a point of inflexion at O so that its expansion at O takes the form

$$y = ax^3 + bx^4 + cx^5 + dx^6 + ex^7 + fx^8 + (9). \tag{4}$$

If a quartic C_4 of the equation (1) has a contact of order 4 with C at O, then

$$\alpha y_1 y_2 + 8(y_1 z_2 - y_2 z_1) + 4\,\frac{b}{a}\,(x_1 y_2 - x_2 y_1) = 0. \tag{5}$$

For a given value α the locus of one of the cusps of C_4 is therefore a line when the other is fixed. Thus, taking α as a parameter, we obtain a pencil of lines

$$\alpha y_1 y + 8(y_1 z - z_1 y) + 4\,\frac{b}{a}\,(x_1 y - y_1 x) = 0, \tag{6}$$

whose center is $(1, 0, b/2a)$, namely, the covariant point O_4 of Bompiani [1].

Suppose that the cuspidal tangent (3) of C_4 passes through a given point $T(x_0, 0, z_0)$ so that

$$-\alpha x_0 y_1 y_2 + \omega_1 x_0 + \omega_3 z_0 = 0.$$

For a given cusp $P_1(x_1, y_1, z_1)$ of C_4 with given cuspidal tangent $P_1 T$

$$-y_1 z_0 x - (x_0 z_1 - x_1 z_0) y + x_0 y_1 z = 0. \tag{7}$$

The locus of the other cusp P_2 is a line

$$9(y_1 z_2 - y_2 z_1) + 4\left(\frac{b}{a} + \frac{z_0}{4x_0}\right)(x_1 y_2 - x_2 y_1) = 0, \tag{8}$$

provided C and C_4 have a contact of the order 4 at O. The lines (7) and (8) coincide if and only if T becomes O_4.

The point of intersection T' of the line (8) and $y = 0$ is

$$\left(1, 0, \frac{4b}{9a} + \frac{z_0}{9x_0}\right) \tag{9}$$

and therefore the correspondence between T and T' is projective with double points at O and O_4. Hence we have the following

THEOREM. *Given a point $P_1(x_1, y_1, z_1)$ outside the inflexional tangent of a curve at O, there is an involution of lines of center P_1 determined by the neighborhood of order 4 of the curve at O, and the double lines pass through O and O_4, respectively.*

3. In order that the curves C and C_4 have a contact of the fifth order we have, besides (5), a further condition, namely,

$$3a\omega_3^2(x_1y_2 + x_2y_1) + 6\omega_1^2y_1y_2 + 6\frac{b}{a}y_1y_2\omega_1\omega_3 + 2\frac{c}{a}y_1y_2\omega_3^2 = 0. \tag{10}$$

If one cusp P_1 of C is fixed, the locus of the other cusp is a nodal cubic having a node at P_1:

$$\begin{aligned}3a(x_1y - y_1x)^2(x_1y + y_1x) + 6y_1y(y_1z - z_1y)^2 \\ + 6\frac{b}{a}y_1y(y_1z - z_1y)(x_1y - y_1x) + \frac{2c}{a}y_1y_2\omega_3^2 = 0.\end{aligned} \tag{11}$$

The two nodal tangents are evidently

$$\frac{x_1y - y_1x}{y_1z - z_1y} = \frac{1}{2cy_1 + 6a^2x_1}[-3by_1 \pm (9b^2y_1^2 - 12acy_1^2 - 36a^3x_1y_1)^{\frac{1}{2}}]. \tag{12}$$

It follows that the curve (11) becomes a cuspidal cubic if and only if P_1 lies on the line $y = 0$ or

$$\left(\frac{b^2}{4a^3} - \frac{c}{3a^2}\right)y - x = 0. \tag{13}$$

The latter is the Bompiani line l_3 (see [1]).

It should be noted that the harmonic conjugate line of OP_1, namely,

$$xy_1 - yx_1 = 0 \tag{14}$$

with respect to the lines (12) is

$$x_1y - y_1x + \frac{2a}{b}(y_1z - z_1y) = 0, \tag{15}$$

and passes through O_4. Hence we arrive at the

THEOREM. *Given a point $P_1(x_1, y_1, z_1)$ outside of l_3 and the inflexional tangent of a curve at its point of inflexion O, we can, by means of the neighborhood of the fifth order of the curve at O, determine a nodal cubic such that P_1 is a node and the harmonic conjugate line of OP_1 with respect to the two nodal tangents passes through the Bompiani osculant O_4.*

On the contrary, if P_1 lies on l_5, then the cubic (11) is a cuspidal cubic with cuspidal tangent:

$$\left(\frac{b^2}{4a^3} - \frac{c}{2a^2}\right)y_1y - y_1x + \frac{2a}{b}(y_1z - z_1y) = 0, \tag{16}$$

which constitutes a pencil of lines with center O_4. Thus we have the

THEOREM. *If one cusp of C_4 is fixed on l_5 and if C_4 has a contact of the fifth order with a curve C at its point of inflexion, then the locus of the other cusp of C_4 is a cuspidal cubic whose cuspidal tangent passes through O_4.*

The line joining $P_1(x_1, y_1, z_1)$ and $(\alpha, 0, \beta)$

$$z = \frac{\beta}{\alpha}x + \left(\frac{z_1}{y_1} - \frac{\beta x_1}{\alpha y_1}\right)y \tag{17}$$

meets the nodal cubic (11) at $P_1(x_1, y_1, z_1)$ and another point

$$\begin{gathered} 3ax_1 + 6\left(\frac{\beta}{\alpha}\right)^2 y_1 - 6\left(\frac{b\beta}{a\alpha}\right)y + \frac{2c}{a}y_1, \\ -3ay_1, \quad \frac{\beta}{\alpha}\left(6ax_1 + 6\frac{\beta^2}{\alpha^2}y_1 - \frac{6b\beta}{a\alpha}y_1 + \frac{2c}{a}y_1\right) - 3az_1. \end{gathered} \tag{18}$$

For a fixed value of β/α the point (18) describes the line

$$3ay_1x + \left(3ax_1 + 6\frac{\beta^2}{\alpha^2}y_1 - 6\frac{b\beta}{a\alpha}y + 2\frac{c}{a}y_1\right)y = 0. \tag{19}$$

Hence there is a projectivity determined between the line (19) and the line

$$x_1y - y_1x = 0. \tag{20}$$

The double elements are found to be $y = 0$ and the line

$$6ax + \left(6\left(\frac{\beta}{\alpha}\right)^2 - 6\frac{b\beta}{a\alpha} + 2\frac{c}{a}\right)y = 0. \tag{21}$$

Thus we obtain a correspondence F_5 between the point $(\alpha, 0, \beta)$ and the line (21), under which O_4 corresponds to l_5. In other words, we have the

THEOREM. *The neighborhood of the fifth order of a curve at its point of inflexion determines a correspondence F_5.*

4. Take the coördinate system such that l_5 is $x = 0$ and O_4, $(1, 0, 0)$. Then

$$b = 0, \qquad c = 0. \tag{22}$$

If C_4 and C have a contact of the sixth order at O, we find, besides (5) and (10), a further condition

(23) $$a^3x_1^2 + (-8dy_1^2 + 16a^2y_1z_1)\frac{\omega_1}{\omega_3} + 18a^2x_1y_1\left(\frac{\omega_1}{\omega_3}\right)^2 + 81ay_1^2\left(\frac{\omega_1}{\omega_3}\right)^4 = 0.$$

From (10) and (23) it follows that for a given point $P_1(x_1, y_1, z_1)$ with $y_1 \neq 0$ there are four related points each of which can be coupled with P_1 as the two cusps of C_4 in order that C_4 and C have a contact of the sixth order at O. If the joins of P_1 with the four related points are harmonic, then the locus of P_1 is easily found to be a cuspidal cubic

(24) $$\frac{3}{2}y(-dy + 2a^2z)^2 - a^5x^3 = 0,$$

whose cusp being $(0, 2a^2, d)$ is the Bompiani osculant O_6 (see [1]). On the contrary, the Bompiani cuspidal cubic [1] is given by the equation

(25) $$y(-dy + 2a^2z)^2 - 4a^5x^3 = 0.$$

Therefore, the Segre invariant of (24) with respect to (25) is 3/8.

Now take the triangle OO_4O_6 for the triangle of reference so that the relations between the new and the old coördinates become

(26) $$x^* = x - \left(\frac{b^2}{4a^3} - \frac{c}{3a^2}\right)y, \qquad y^* = y,$$
$$z^* = \frac{1}{2a^2}\left[2a^2z - abx - \left(d - \frac{2bc}{a} + \frac{b^3}{a^2}\right)y\right];$$

the expansion of the curve (4) then takes the form

(27) $$y = ax^3 + px^7 + qx^8 + (9),$$

where

(28) $$p = e - \frac{5}{2}\frac{bd}{a} + 5\frac{b^2c}{a^2} - \frac{33}{16}\frac{b^4}{a^3} - \frac{4}{3}\frac{c^2}{a},$$
$$q = f - 3\frac{be}{a} + 6\frac{b^2d}{a^2} - \frac{39}{4}\frac{b^3c}{a^3} + \frac{145}{32}\frac{b^5}{a^4} + \frac{19}{3}\frac{bc^2}{a^2}$$
$$- 3\frac{d}{a^2} + \frac{b^4}{8a^3} - 3\frac{b^3d}{2a^3} + 3\frac{bcd}{a^2}.$$

In the new coördinate system every line of the pencil with center $(\alpha, 0, \beta)$, where $\alpha\cdot\beta \neq 0$, in general, possesses a pair of points which can be taken for the two cusps of C_4 having at O a contact of the sixth order with C. When the line varies in the pencil, the locus of the point is, after (23), a conic:

(29) $$a^3x^2 - 16a^2\left(\frac{\beta}{\alpha}\right)yz + 18a^2\left(\frac{\beta}{\alpha}\right)xy + 81ay^2\left(\frac{\beta}{\alpha}\right)^4 = 0,$$

which touches $y = 0$ at $(0, 0, 1)$. Let P^* be the pole of l_5 with respect to the conic (29); then

$$(OO_4\,,\,P^*P) = \frac{9}{8}. \tag{30}$$

If we take β/α as a parameter, the conic (29) envelopes a cuspidal cubic

$$-3z^2y + ax^3 = 0. \tag{31}$$

The latter has at O the Segre invariant $1/3$ with the Bompiani cubic.

Similarly, the polar line of O_6

$$-16az + 18a\frac{\beta}{\alpha}x + 16z\left(\frac{\beta}{\alpha}\right)^3 z = 0 \tag{32}$$

also envelopes a cubic

$$4yz^2 + \frac{1}{12}az^3 = 0, \tag{33}$$

the Segre invariant in this case being $1/48$.

In general, for a point $P_1(x_1\,,\,y_1\,,\,z_1)$ $(y_1 \neq 0)$ we can determine four related points by the neighborhood of the sixth order at the point of inflexion, and accordingly obtain a conic through P_1 and its four related points, namely,

$$\frac{28}{9}a^3x_1y_1^3x^2 + \frac{16a^3x_1^2y_1^2}{81}xy + \left(-\frac{196}{81}a^2y_1^2z_1^2 - \frac{268}{81}a^3y_1x_1^3\right) y^2 + \frac{520}{81}a^2y_1^3z_1yz - 4c^2y_1^4z^2 = 0. \tag{34}$$

This conic passes through $P_1(x_1\,,\,y_1\,,\,z_1)$ and cuts $y = 0$ in the two harmonic points of O and O_4 :

$$\left(1, 0, \pm\left(\frac{9y_1}{7ax_1}\right)^{\frac{1}{2}}\right).$$

Since the polar line of O_4 with respect to the conic (34) is

$$7y_1x + \frac{2}{9}x_1y = 0, \tag{35}$$

we have that the cross-ratio of the four lines (35), OP_1, l_5, and the inflexional tangent is $-2/63$. As to the polar line of $O(0, 0, 1)$ with respect to the conic (34), namely,

$$\frac{130}{81}z_1y - y_1z = 0, \tag{36}$$

we have that the cross-ratio of the four lines (36), O_4P_1, the inflexional tangent at O of C, and the cuspidal tangent of the Bompiani cubic is $81/130$.

In virtue of (29) and (34) a correspondence between points and conics is obtained. Denoting it by F_6, we have the

THEOREM. *The neighborhood of the sixth order at a point of inflexion of a curve determines a covariant point O_6 and a correspondence F_6.*

It should be noted that if the conic (34) degenerates into lines, the locus of P_1 is

$$1792yz^2 - 36765ax^3 = 0, \tag{37}$$

whose Segre invariant with the Bompiani cubic is 36765/1792.

5. If C_4 has at O a contact of the seventh order with C a further condition is required, namely,

$$7a^2x_1^2\omega_3^2 + \frac{9p}{a}y_1^2\omega_3^2 - 50ay_1z_1\omega_1\omega_3 - 36ax_1y_1\omega_1^2 = 0. \tag{38}$$

The locus of the cusps of C_4 in consideration can be found by eliminating ω_1/ω_3 from (23) and (38). Putting

$$\frac{\omega_1}{\omega_3} = \tau,$$

we know that the locus is given by

$$7a^2x^2 + 9\frac{p}{a}y^2 - 50ayz\tau - 36axy\tau^2 = 0,$$

$$a^3x^2 + 16a^2yz\tau + 18a^2xy\tau^2 + 81ay^2\tau^4 = 0,$$

namely, it is of the parametric equations

$$\begin{aligned}\frac{x}{y} &= \frac{1}{a}\left(-\tau^2 \pm \left(-24\tau^4 - \frac{8p}{9a}\right)^{\frac{1}{2}}\right),\\ \frac{z}{y} &= \frac{-1}{2a\tau}\left(5\tau^4 - \frac{p}{9a} \pm 2\tau^2\left(-24\tau^4 - \frac{8p}{9a}\right)^{\frac{1}{2}}\right).\end{aligned} \tag{39}$$

For a fixed value of τ ($\neq 0$) there are two corresponding points (39). This pair of points can be taken for the two cusps of C_4 such that C_4 at O has a contact of seventh order with C. On the same tangent $y = 0$ there are, however, four points, other than O, through each of which no line can be drawn such that a suitable pair of points on it becomes the cusps of the quartic C_4 in consideration. They are, in fact, of the coördinates

$$(1, 0, \tau_i) \qquad (i = 1, 2, 3, 4),$$

where τ_i are the roots of

$$\tau^4 + \frac{p}{27a} = 0. \tag{40}$$

For these τ_i $(i = 1, 2, 3, 4)$ the corresponding points (39) are

$$\begin{aligned} \frac{x}{y} &= -\frac{1}{a}\left(\cos\left(k + \frac{\epsilon}{2}\right)\pi + i\sin\left(k + \frac{\epsilon}{2}\right)\pi\right)\left|\frac{-p}{27a}\right|^{1/4}, \\ \frac{z}{y} &= -\frac{1}{a}\left(\cos\left(\frac{6k + 3\epsilon}{4}\right)\pi + i\sin\left(\frac{6k + 3\epsilon}{4}\right)\pi\right)\left|\frac{-p}{27a}\right|^{3/4} \end{aligned} \tag{41}$$

$$(k = 1, 2, 3, 4),$$

where $\epsilon = 0$ or 1 according as $-p/a > 0$ or < 0. In the pencil of conics determined by the four points (41) there is a particular one not only passing through O_4 and O_6 but also tangent to $y = 0$ and $x = 0$ at O_4 and O_6, respectively. This conic is

$$z^2 = \frac{16p}{27a^2}xy. \tag{42}$$

If $e \neq 0$ we take a point of intersection of the conic (42) and the Bompiani cubic

$$z^2y - ax^3 = 0 \tag{43}$$

for the unit point such that

$$a = 1, \qquad p = \frac{27}{16}, \tag{44}$$

and, in consequence, the canonical expansion of the curve C at O becomes

$$y = x^3 + \frac{27}{16}x^7 + qx^8 + (9). \tag{45}$$

As the line through the two points (39) is

$$\tau^2x + \frac{1}{2a}\left(7\tau^4 - \frac{p}{9a}\right)y + \tau z = 0, \tag{46}$$

it takes the limiting position

$$\tau_i^2x + \frac{1}{2a}\left(7\tau_i^4 - \frac{p}{9a}\right)y + \tau_i z = 0 \qquad (i = 1, 2, 3, 4) \tag{47}$$

when $\tau \to \tau_i$ $(i = 1, 2, 3, 4)$, where τ_i are the roots of (30). There is a pencil of conics determined by the four lines (47). If we take (44) for unit point, the equation of this pencil is

$$v^2 + 25u^2 + \kappa(16uv + 5w^2) = 0, \tag{48}$$

where κ is a parameter. If the conic (48) touches the line

$$x + y + z = 0, \tag{49}$$

it must be of the equation

$$v^2 + 25u^2 - \frac{26}{21}(16uw + 5w^2) = 0, \tag{50}$$

provided OO_4O_6 is, as before, the triangle of reference and (44) the unit point. A point of inflexion of a plane curve is said to be of the first class when in (27)

$$ap \neq 0, \tag{51}$$

the second class when

$$ap = 0, \qquad aq \neq 0 \tag{52}$$

and the third class when

$$p = q = 0.$$

Of course, *the canonical expansion* (45) *is not valid for the points of inflexion of the second or the third class.*

6. If C_4 has at O a contact of the eighth order with C there is one more required condition

$$12axy\tau^3 + 2ayz\tau^2 + 5a^2x^2\tau + 3\frac{p}{a}y^2\tau + \frac{q}{a}y^2 = 0. \tag{53}$$

Substituting (39) in (53) we have

$$132\tau^5 + \frac{4p}{3a}\tau - \frac{q}{a} = 0. \tag{54}$$

By (54) we can find τ_i $(i = 1, \cdots, 5)$ and then obtain five pairs of points (39) such that each pair of them can be taken for the cusps of C_4 in consideration. If the point of inflexion is of the second class, that is, $p = 0$ and $q \neq 0$, the five roots of (54) are

$$\tau_k = \epsilon^k \left|\frac{q}{132a}\right|^{1/5}, \quad \epsilon^5 = \operatorname{sgn}\left(\frac{-q}{a}\right) \qquad (\epsilon \neq \pm 1; k = 1, \cdots, 5). \tag{55}$$

In this case the five lines determined by each pair of points (39)

$$z + \frac{3}{a}\tau_k^3 y + \frac{\tau_k}{2}x = 0 \qquad (k = 1, \cdots, 5) \tag{56}$$

are tangent to the conic

$$\frac{2q}{132a}uw - \frac{a^2}{9}v^2 = 0, \tag{57}$$

that is,

$$3qy^2 - 88a^3xz = 0. \tag{58}$$

Now we can take a point of intersection of the Bompiani cubic (43) and the conic (58) for the unit point such that

$$a = 1, \qquad q = \frac{88}{3},$$

and the canonical expansion of the curve at its point of inflexion of the second class is

$$y = x^3 + \frac{88}{3}x^8 + (9).$$

If $p \neq 0$, that is, the point of inflexion is of the first class, we have from (39) five pairs of roots of

$$132\tau^5 + \frac{9}{4}\tau - q = 0. \tag{59}$$

The line determined by every pair of points (59) is

$$\tau_i^2 x + y\left(\frac{7}{2}\tau_i^4 - \frac{3}{32}\right) + \tau_i z = 0 \qquad (i = 1, \cdots, 5), \tag{60}$$

and the five lines thus obtained further determine a conic, namely,

$$v^2 - \frac{49q}{528}uw + \frac{7\cdot 97}{32\cdot 22}u^2 - \frac{11\cdot 3^3}{7\cdot 2^7 q}vw - \frac{11\cdot 3^7}{7^2\cdot 2^8 q^2}uv - \frac{3^{10}}{7^2\cdot 2^{12}\cdot q^2}u^2 = 0. \tag{61}$$

The pole of the inflexional tangent of C at O, namely, $u = w = 0$, with respect to the curve (61) is

$$P^*\left(-\frac{11\cdot 3^7}{7^2\cdot 2^8\cdot q^2},\ 2,\ -\frac{11\cdot 3^3}{7\cdot 2^7 q}\right). \tag{62}$$

Since the determination of the unit point is geometrical, we may express in terms of the double ratio D of $z = 0$, $y = 0$, $z - y = 0$ and P^*O_4 :

$$q = -\frac{11\cdot 3^3}{2^8\cdot 7}D.$$

BIBLIOGRAPHY

1. E. BOMPIANI, *Per lo studio proiettivo differenziale delle singolarità*, Boll. Dell'Unione Mat. Ital., vol. 5(1926), pp. 118–120.
2. SU-CHENG CHANG, *On the cusp of a plane curve.*
3. B. SU, *Note on the projective geometry of singularities of plane curves*, Journal of the Chinese Mathematical Society, vol. 2(1940), pp. 153–173.

NATIONAL UNIVERSITY OF CHEKIANG, MEITAN, KWEICHOW, CHINA.

THE SINGULARITY S_1^m OF A PLANE CURVE

By Su-Cheng Chang

1. By a singularity S_1^m of a plane curve we mean the point at which the tangent to the curve has a contact of order m with the curve. If this point is taken for the origin $O(0, 0, 1)$ and the tangent for $y = 0$, then the expansions of the curve in the neighborhood of O become

$$(1) \qquad x = s \sum_0^\infty a_r s^r, \qquad y = s^m \sum_0^\infty b_r s^r, \qquad z = 1 + \sum_0^\infty c_r s^r, \qquad a_0 b_0 \neq 0.$$

In particular, when $m = 3$, the singularity which is a point of inflexion has been studied by E. Bompiani [1] and the author [2]. It is B. Su [4], [5] who generalizes Bompiani's osculants to a curve with a representable singularity of high order. In a recent paper [3] we have studied the singularity S_1^4 in detail and obtain the canonical expansions of two species of S_1^4 that had been classified projectively.

It is natural to extend our method of representing the neighborhood of various orders of an S_1^4 to the study of an S_1^m $(m > 4)$. Here we investigate only the representable singularity considered by Su, namely, the singularity for which the invariant point O_{2m} exists, and give a geometrical interpretation of the conditions for a representable S_1^m, as these have been derived analytically by Su.

There are other covariant figures, besides O_{m+1}, l_{2m-1} and O_{2m}, determined by the neighborhoods of high orders of a representable S_1^m. A formulation of these elements as well as a supplement to the canonical expansion of Su for two species of a representable S_1^m forms the main object of this note.

2. Suppose that a curve C has a singular point S_1^m at $O(0, 0, 1)$, so that the expansion can be written in the form (1). In what follows we shall utilize an algebraic curve C_m of order m having a node, a singular point S_1^m, and an $(m - 2)$-ple point with coincident tangent. Let $P_1(x_1, y_1, z_1)$ and $P_2(x_2, y_2, z_2)$ be the $(m - 2)$-ple point and the node of C_m, respectively, and let

$$\frac{x_1 y - y_1 x}{y_1} + \rho \frac{\omega_1 x + \omega_2 y + \omega_3 z}{y_1 y_2} = 0$$

be the equation of the common tangent of C_m at P_1; the equation of C_m, which has a contact of order m with C at O, can be written as

$$y^2 \left(\frac{x_1 y - y_1 x}{y_1} + \rho \frac{\omega_1 x + \omega_2 y + \omega_3 z}{y_1 y_2} \right)^{m-2} - 2 \frac{y_1 y_2}{\omega_3} y \left(\frac{x_1 y - y_1 x}{y_1} \right)^{m-1}$$

Received April 22, 1942.

$$
\begin{aligned}
(2)\qquad & + \left(\frac{y_1y_2}{\omega_3}\right)^2 \left(\frac{x_1y - y_1x}{y_1}\right)^m \\
& + y \sum_1^{m-2} A_i \left(\frac{x_1y - y_1x}{y_1}\right)^{m-i-1} \left(\frac{\omega_1x + \omega_2y + \omega_3z}{y_1y_2}\right)^i \\
& + (-1)^{m+1} \frac{a_0^m}{b_0} \left(\frac{y_1y_2}{\omega_3}\right)^{m+1} \left(\frac{\omega_1x + \omega_2y + \omega_3z}{y_1y_2}\right)^{m-1} y = 0,
\end{aligned}
$$

where

$$\omega_1 = y_1z_2 - y_2z_1, \qquad \omega_2 = z_1x_2 - z_2x_1, \qquad \omega_3 = x_1y_2 - x_2y_1,$$

$$\rho = \frac{-1}{m-2} \frac{\omega_3}{y_1y_2} A_1$$

and A_i $(i = 1, \cdots, m-2)$ are arbitrary constants. In order that C_m have at O a contact of order $m + 1$ with C it is necessary and sufficient that

$$
\begin{aligned}
(3)\qquad & -b_0a_0A_{m-2}\left(\frac{\omega_3}{y_1y_2}\right)^{m-2} + (-1)^m \left(\frac{y_1y_2}{\omega_3}\right)^2 ma_0^{m-1}a_1 + (-1)^{m+1} \frac{a_0^m}{b_0} \left(\frac{y_1y_2}{\omega_3}\right)^{1+m} \\
& \times \left[\left(\frac{\omega_3}{y_1y_2}\right)^{m-1} b_1 + (m-1)\left(\frac{\omega_3}{y_1y_2}\right)^{m-1} b_0c_1 + \frac{(m-1)\omega_1\omega_3^{m-2}}{(y_1y_2)^{m-1}} a_0b_0\right] = 0.
\end{aligned}
$$

The quantity A_{m-2} thus determined is called the $(m + 1)$-th parameter [3]. For given point P_1 and fixed A_{m-2} the locus of P_2 is

$$
\begin{aligned}
& -a_0b_0A_{m-2}(x_1y - y_1x)^{m+1} + (-1)^{m+1}(m-1)a_0^{m+1}(y_1y)^m(y_1z - z_1y) \\
(4)\qquad & + (-1)^{m+1} \frac{a_0^m}{b_0} \Big[b_1(y_1y)^m(x_1y - y_1x) + (m-1)b_0c_1(y_1y)^m(x_1y - y_1x) \\
& \qquad - m \frac{a_1b_0}{a_0} (y_1y)^m(x_1y - y_1x) \Big] = 0.
\end{aligned}
$$

As A_{m-2} varies the curves (4) always pass through P_1 and constitute a pencil, called the pencil of the $(m + 1)$-th curves with center P_1. The tangents of these curves at their center coincide in the line

$$(5)\qquad \left(b_1 + \overline{m-1}b_0c_1 - m\frac{a_1b_0}{a_0}\right)(x_1y - y_1x) + (m-1)a_0b_0(y_1z - z_1y) = 0,$$

which clearly passes through the point

$$(6)\qquad \left((m-1)a_0b_0,\ 0,\ b_1 + (m-1)b_0c_1 - m\frac{a_1b_0}{a_0}\right)$$

which is independent of P_1. Denoting, after Su [4], [5], this point by O_{m+1}, we have the

THEOREM. *Given a point P_1 outside of the tangent of a plane curve at its S_1^m, there is determined a pencil of the $(m+1)$-th curves with center P_1. The curves of this pencil have at P_1 a common tangent which passes through the invariant point O_{m+1}.*

3. Put

$$x^r = s^r\Big(\sum_0^\infty a_\nu s^\nu\Big)^r = s^r \sum_0^\infty A_{r,\mu} s^\mu,$$

$$z^r = \Big(1 + \sum_1^\infty c_\nu s^\nu\Big)^r = \sum_0^\infty C_{r,\mu} s^\mu,$$

$$U_{1r,k} = \sum_{\substack{\nu,\mu,\tau=0\\ \nu+\mu+\tau=k}}^{k} b_\nu A_{r,\mu} C_{m-r-1,\tau},$$

so that

$$yx^r z^{m-r-1} = s^{m+r} \sum_0^\infty U_{1r,\tau} s^\tau.$$

If O_{m+1} is taken for (1, 0, 0), then

$$b_1 + (m-1)b_0c_1 - m\,\frac{a_1b_0}{a_0} = 0$$

and

$$A_{m-2} = (-1)^{m+1}(m-1)\,\frac{a_0^m}{b_0}\cdot\frac{(y_1y_2)^m\omega_1}{(\omega_3)^{m+1}}. \tag{7}$$

In order that C_m have a contact of order $m+2$ with C another condition is required, namely,

$$\begin{aligned} &A_{m-3}b_0a_0^2\omega_3^{m+1} + (-1)^{m+1}\frac{a_0^m}{b_0}U_{10,2}(y_1y_2)^{m-1}\omega_3^2 \\ &\quad - (-1)^{m+1}\cdot\frac{(m-1)(m-2)}{2}\,\frac{a_0^m}{b_0}U_{12,0}(y_1y_2)^{m-1}\omega_1^2 \\ &\qquad + (-1)^m A_{m,2}(y_1y_2)^{m-1}\omega_3^2 = 0. \end{aligned} \tag{8}$$

By means of (8) we can choose A_{m-3} for prescribed P_1 and P_2 in order to obtain the C_m under consideration. We call A_{m-3} determined by (8) the $(m+2)$-th parameter of P_1 and P_2.

When P_1 and A_{m-3} are fixed, the locus of P_2 is

$$
\begin{aligned}
(9)\qquad & Cb_0a_0^2(x_1y - y_1x)^{m+1} + (-1)^{m+1}\frac{a_0^m}{b_0}U_{10,2}(y_1y)^{m-1}(x_1y - y_1x)^2 \\
& + (-1)^m\begin{bmatrix} m-1 \\ 2 \end{bmatrix}a_0^{m+2}(y_1y)^{m-1}(y_1z - z_1y)^2 \\
& + (-1)^m(y_1y)^{m-1}A_{m,2}(x_1y - y_1x)^2 = 0,
\end{aligned}
$$

where C is an arbitrary constant. We call the curves (9) the $(m + 2)$-th curves with center P_1. Since P_1 is the node of these curves, the two nodal tangents at P_1 which are independent of the $(m + 2)$-th parameter are given by

$$
(10)\qquad \left(\frac{a_0^m}{b_0}U_{10,2} - A_{m,2}\right)(x_1y - y_1x)^2 - \begin{bmatrix} m-1 \\ 2 \end{bmatrix}a_0^{m+2}(y_1z - z_1y)^2 = 0.
$$

These lines intersect the tangent $y = 0$ at two invariant points

$$
(11)\qquad \left(\begin{bmatrix} m-1 \\ 2 \end{bmatrix}^{\frac{1}{2}}\cdot a_0^{(m+2)/2},\ 0,\ \pm\left[\frac{a_0^m}{b_0}U_{10,2} - A_{m,2}\right]^{\frac{1}{2}}\right).
$$

It is clear that these points, denoted by $O_{m+2,1}$ and $O_{m+2,2}$, coincide if and only if

$$
(12)\qquad \frac{a_0^m}{b_0}U_{10,2} - A_{m,2} = 0,
$$

which is nothing but the first of the conditions in order that the singularity S_1^m be representable. (See [3], [4], [5]. We call these conditions the conditions of Su.) Hence we have the

THEOREM. *The neighborhood of order $m + 2$ of a curve C at its singular point S_1^m determines a three parameter system of the $(m + 2)$-th curves, whose nodal tangents at their center pass through two covariant points $O_{m+2,1}$ and $O_{m+2,2}$ at the tangent of C at S_1^m. These points coincide if and only if the first condition of Su is satisfied.*

If

$$
(13)\qquad \frac{a_0^m}{b_0}U_{10,k} - A_{m,k} = 0 \qquad (k = 2, \cdots, r - 1),
$$

for $3 \leq r < m - 1$, C_m has at O a contact of order $m + r$ with C if and only if

$$
\begin{aligned}
& \omega_3^{m+1}A_{m-l} - (-1)^{m+1}\begin{bmatrix} m-1 \\ l-1 \end{bmatrix}\frac{a_0^m}{b_0}(y_1y_2)^{m-l+2}\omega_1^{l-1} = 0 \quad (l = 2, \cdots, r), \\
(14)\qquad & \omega_3^{m+1}A_{m-r-1} - (-1)^{m+1}\begin{bmatrix} m-1 \\ r \end{bmatrix}\frac{a_0^m}{b_0}(y_1y_2)^{m-r+1}\omega_1^r \\
& + \frac{1}{U_{1r,0}}(-1)^{r+m+1}\omega_3^r(y_1y_2)^{m-r+1}\left(A_{m,r} - \frac{a_0^m}{b_0}U_{10,r}\right) = 0.
\end{aligned}
$$

As before, we can similarly define the $(m + r)$-th parameter, the $(m + r)$-th curve with center P_1 and the r invariant points

$$(15) \qquad \left(\left(\begin{bmatrix} m-1 \\ r \end{bmatrix} a_0^{m+r}\right)^{1/r}, 0, \left[(-1)^{r-1}\left(A_{m,r} - \frac{a_0^m}{b_0} U_{10,r}\right)\right]^{1/r} \epsilon^i\right)$$

$$(i = 1, \cdots, r, \epsilon^r = 1, \epsilon \neq 1),$$

the latter being denoted by $O_{m+r,i}$ $(i = 1, \cdots, r)$. If two of them coincide, all of them necessarily coincide with O_{m+1}, since in this case

$$(16) \qquad A_{m,r} - \frac{a_0^m}{b_0} U_{10,r} = 0,$$

which is one of the conditions of Su.

Hence we have the

THEOREM. *Suppose the first $r - 2$ conditions of Su are satisfied at a singularity S_1^m of a curve, namely,*

$$\frac{a_0^m}{b_0} U_{10,k} - A_{m,k} = 0 \qquad (k = 2, \cdots, r - 1; 2 < r < m - 1).$$

Then there exist the $(m + r)$-th curves determined by the neighborhood of order $m + r$ of the curve at the singularity S_1^m. The center of these curves is an r-ple point and each of the r tangents of the curve passes through an invariant point $O_{m+r,i}$ $(i = 1, \cdots, r)$. If two of these points coincide, then all of them are coincident with O_{m+1}. The necessary and sufficient condition for this is that the $(r - 1)$-th condition among the conditions of Su be satisfied, namely,

$$\frac{a_0^m}{b_0} U_{10,r} - A_{m,r} = 0.$$

4. In what follows we consider only the case where O is a representable singularity S_1^m, namely,

$$(17) \qquad \frac{a_0^m}{b_0} U_{10,r} - A_{m,r} = 0 \qquad (r = 2, \cdots, m - 2).$$

Such a singularity is denoted by $S_{1,m-3}^m$. If C_m has at O a contact of order $2m - 1$ with C, then we have

$$(18) \qquad \omega_3^{m+1} A_{m-j} - (-1)^{m+1} \begin{bmatrix} m-1 \\ j-1 \end{bmatrix} \frac{a_0^m}{b_0} (y_1 y_2)^{m-j+2} \omega_1^{j-1} = 0$$

$$(j = 1, \cdots, m - 2)$$

and a further condition

$$(19)\qquad \begin{aligned} 2b_0a_0^{m-1} - \frac{a_0^m}{b_0} U_{10,m-1} \frac{y_1y_2}{\omega_3} + (-1)^{m+1}a_0^{2m-1} \frac{y_1y_2\omega_1^{m-1}}{\omega_3^m} \\ + \frac{y_1y_2}{\omega_3}\left[A_{m,m-1} - m\frac{x_1}{y_1} b_0a_0^{m-1}\right] = 0. \end{aligned}$$

When P_1 is fixed outside $y = 0$ the locus of $P_2(x_2, y_2, z_2)$ $(y_2 \neq 0)$ is a curve

$$(20)\qquad \begin{aligned} 2b_0a_0^{m-1}(x_1y - y_1x)^m - \frac{a_0^m}{b_0} U_{10,m-1}y_1y(x_1y - y_1x)^{m-1} \\ + (-1)^{m+1}a_0^{2m-1}y_1y(y_1z - z_1y)^{m-1} \\ + y_1y(x_1y - y_1x)^{m-1}\left[A_{m,m-1} - \frac{mx_1}{y_1} b_0a_0^{m-1}\right] = 0. \end{aligned}$$

Obviously, P_1 is an $(m - 1)$-ple point of (20). A necessary and sufficient condition in order that the tangents at P_1 be one and the same is that

$$(21)\qquad \frac{a_0^m}{b_0} U_{10,m-1} - A_{m,m-1} + m\frac{x_1}{y_1} a_0^{m-1}b_0 = 0.$$

That is, P_1 must be a point on the line

$$(22)\qquad y\left(\frac{a_0^m}{b_0} U_{10,m-1} - A_{m,m-1}\right) + ma_0^{m-1}b_0x_1 = 0,$$

which is the line l_{2m-1} (see [4], [5]). Thus we arrive at the

THEOREM. *If an arbitrarily assigned point outside the tangent at an $S_{1,m-3}^m$ of a curve C is taken for P_1 of C_m and if C_m has at the $S_{1,m-3}^m$ a contact of order $2m - 1$ with C, then the locus of P_2 is a curve of order m having P_1 for an $(m - 1)$-ple point. The $(m - 1)$ tangents at P_1 become one and the same if and only if P_1 lies on the line l_{2m-1}.*

Taking now a point $T(\lambda, 0, -1)$ on $y = 0$ we can obtain a line P_1T which intersects (20) at P_1 and another point Q on the following line

$$(23)\qquad \begin{aligned} 2b_0a_0^{m-1}(x_1y - y_1x) \\ - y_1y\left\{\frac{a_0^m}{b_0} U_{10,m-1} - (-1)^{m+1}\frac{a_0^{2m-1}}{\lambda^{m-1}} - A_{m,m-1} + m\frac{x_1}{y_1} b_0a_0^{m-1}\right\} = 0. \end{aligned}$$

When P_1 varies on a line through O,

$$(24)\qquad y - \mu x = 0,$$

the locus of Q is found to be

$$(25)\quad \begin{aligned} &2b_0a_0^{m-1}(y - \mu x) \\ &- \left[\frac{a_0^m}{b_0} U_{10,m-1}\mu - (-1)^{m+1}\frac{a_0^{2m-1}}{\lambda^{m-1}}\mu - \mu A_{m,m-1} + mb_0a_0^{m-1}\right] y = 0. \end{aligned}$$

For a fixed λ the correspondence between (24) and (25) is a projectivity whose double elements are $y = 0$ and

$$(26)\quad \frac{a_0^m}{b_0} U_{10,m-1} + (-1)^m \cdot \frac{a_0^{2m-1}}{\lambda^{m-1}} - A_{m,m-1} + mb_0a_0^{m-1}\frac{x}{y} = 0.$$

Denoting the correspondence between the point T and the line (26) by F_{2m-1} and remarking that the corresponding line of O_{m+1} is precisely l_{2m-1}, we obtain the

THEOREM. *A correspondence F_{2m-1} is determined by the neighborhood of order $2m - 1$ of a curve at its $S_{1,m-3}^m$ such that the range of points on the tangent at $S_{1,m-3}^m$ of the curve corresponds to the pencil of lines with center $S_{1,m-3}^m$. In particular, O_{m+1} corresponds to l_{2m-1} under F_{2m-1}.*

5. Now take the line l_{2m-1} for $x = 0$, so that

$$\frac{a_0^m}{b_0} U_{10,m-1} - A_{m,m-1} = 0,$$

and (19) becomes

$$(27)\quad \begin{aligned} &2b_0(x_1y_2 - x_2y_1)^m + (-1)^{m+1}a_0^m(y_1z_2 - y_2z_1)^{m-1}y_1y_2 \\ &\qquad - mb_0x_1y_2(x_1y_2 - x_2y_1)^{m-1} = 0. \end{aligned}$$

If C_m has at $S_{1,m-3}^m$ a contact of order $2m$ with C, then the points P_1 and P_2 are related by (27) and a further condition

$$(28)\quad \begin{aligned} &\left(\frac{m-1}{m-2}\right)^{m-2} \omega_1(2\omega_3 - mx_1y_2)^{m-3} + \frac{1}{a_0^m b_0}\left(A_{m,m} - \frac{a_0^m}{b_0} U_{10,m}\right) y_1y_2\omega_3^{m-3} \\ &\qquad - (m-1)y_2\omega_3^{m-4}(x_1\omega_1 + y_1\omega_2) = 0. \end{aligned}$$

For a fixed P_1, (27) and (28) represent two curves having P_1 and $(m - 2)^2$ points $B_j(j = 1, \cdots, \overline{m - 2}^2)$ outside $y = 0$ in common. Each of them may be taken for P_2 so that the corresponding C_m has at $S_{1,m-3}^m$ a contact of order $2m$ with C. The curve given by the equation

$$(29)\quad \begin{aligned} &\left(\frac{m-1}{m-2}\right)^{m-2}(y_1z - z_1y)[(2 - m)x_1y_1 - 2y_1x]^{m-3} \\ &\qquad - (m-1)x_1y(y_1z - z_1y)(x_1y - y_1x)^{m-4} \\ &\qquad + \frac{1}{a_0^m b_0}\left(A_{m,m} - \frac{a_0^m}{b_0} U_{10,m}\right) y_1y(x_1y - y_1x)^{m-3} \\ &\qquad - (m-1)y_1y(z_1x - x_1z)(x_1y - y_1x)^{m-4} = 0 \end{aligned}$$

passes through P_1, O_{m+1}, B_j $(j = 1, \cdots, (m-2)^2)$ and has O for an $(m-3)$-ple point with coincident tangents. Since x_1/y_1 and z_1/y_1 can vary, the curves (29) constitute a two parameter system which represents the neighborhood of order $2m$ of the curve at the $S_{1,m-3}^m$. The polar conic of O_{m+1} with respect to (29) is

$$\left(\frac{m-1}{m-2}\right)^{m-2} \cdot 2^{m-3}(y_1z - z_1y)(x_1y - y_1x)$$
$$(30) \qquad - 2^{m-4} \cdot m\left(\frac{m-1}{m-2}\right)^{m-2} x_1y(y_1z - z_1y)$$
$$+ \left[\frac{y_1}{a_0^m b_0}\left(A_{m,m} - \frac{a_0^m}{b_0} U_{10,m}\right) + (m-1)z_1\right] y(x_1y - y_1x) = 0,$$

which gives a correspondence F_{2m} between a point P_1 outside $y = 0$ and a conic (30). The latter degenerates if and only if P_1 lies on the lines $x = 0$ and

$$(31) \qquad y\left(A_{m,m} - \frac{a_0^m}{b_0} U_{10,m}\right) + (m-1)a_0^m b_0 z = 0$$

which intersect at the point O_{2m}

$$(32) \qquad \left(0, (m-1)a_0^m b_0, \frac{a_0^m}{b_0} U_{10,m} - A_{m,m}\right).$$

Thus we have the

THEOREM. *A correspondence F_{2m} is determined by the neighborhood of order $2m$ of a curve at its $S_{1,m-3}^m$, such that a point not lying on the tangent of the curve at $S_{1,m-3}^m$ corresponds to a conic. The point O_{2m} is obtained as the intersection of the two lines which stand for the locus of the points that correspond to degenerate conics under F_{2m}.*

6. If the point O_{2m} is taken for the remaining vertex of the triangle of reference, then

$$\frac{a_0^m}{b_0} U_{10,m} - A_{m,m} = 0.$$

In order that C_m have a contact of order $2m$ with C at $S_{1,m-3}^m$ it is necessary and sufficient that the equations (18), (27),

$$(33) \qquad \left(\frac{m-1}{m-2}\right)^{m-2} \omega_1(2\omega_3 - mx_1y_2)^{m-3} + (m-1)y_2z_1\omega_3^{m-3} = 0,$$

and

$$(34) \qquad U_{21,0}\left[(m-2)^2 \frac{y_2z_1\omega_3^{m-3}}{\omega_1^{m-2}} + (m-1)(m-2)(-1)^m \frac{a_0^m}{b_0}\frac{\omega_1^2}{\omega_3^4} x_1y_1y_2^2\right]$$
$$+ (-1)^m\left(\frac{y_1y_2}{\omega_3}\right)^2\left(A_{m,m+1} - \frac{a_0^m}{b_0} U_{10,m+1}\right) = 0$$

hold.

To classify the $(m-2)$-ple points P_1 of all curves C_m which have at the $S^m_{1,m-3}$ a contact of order $2m+1$ with C, let us take the point where P_1P_2 meets the tangent $y = 0$. If this is fixed, then

$$\frac{\omega_3}{\omega_1} = \rho,$$

where ρ is an arbitrary constant. From (27), (33), and (34), it follows that any one of the $m-3$ points of intersection R_i $(i = 1, \cdots, m-3)$ outside $y = 0$ of the curves

$$z^{m-5}\left(\frac{m-1}{m-2}\right)^{m-2}\left(\frac{(m-2)^2a_0^{m+1}b_0\rho}{yS_{m+1}\rho^2 + (m-1)(m-2)a_0^{m+1}b_0x}\right)^{m-4} + (m-1)\frac{b_0}{a_0^m}\frac{\rho^m}{y} = 0 \tag{35}$$

and

$$2[y^2S_{m+1}\rho^2 + (m-1)(m-2)b_0a_0^{m+1}xy] + (-1)^m(m-2)^2a_0b_0^2z\left(mx\rho^m + (-1)^m\frac{a_0^m}{b_0}\rho y\right) = 0, \tag{36}$$

where

$$S_{m+1} = A_{m,m+1} - \frac{a_0^m}{b_0}U_{10,m+1},$$

can be taken for the $(m-2)$-ple point of the C_m in consideration. The coördinates of the node of this C_m may be determined by $\omega_3/\omega_1 = \rho$ and (34), provided P_1 is given. In this case P_1P_2 passes through the point $(-\rho, 0, 1)$. The conic (36) passes through O, O_{m+1} and the $m-3$ points R_i $(i = 1, \cdots, m-3)$ and the tangent at O is the corresponding line

$$mx\rho^{m-1}\,(-1)^m\frac{a_0^m}{b_0}y = 0 \tag{37}$$

of $P_\rho(\rho, 0, 1)$ under F_{2m-1}. Moreover, the polar line of O_{2m} with respect to this conic is

$$2[2yS_{m+1}\rho^3 + (m-1)(m-2)b_0a_0^{m+1}x] + (m-2)^2a_0^{m+1}b_0\rho z = 0, \tag{38}$$

which intersects $y = 0$ at the point $P'_\rho((m-2)\rho, 0, -2(m-1))$. The cross ratio of P_ρ, P'_ρ, O and O_{m+1} always equals $2(m-1)/(m-2)$. Thus by means of the neighborhood of order $2m+1$ of the curve at its $S^m_{1,m-3}$ we can define a correspondence F_{2m+1} between a point on $y = 0$ and a conic. Since ρ is a parameter, the conics (36) constitute a system. It is easily seen that the line (38) envelopes a conic with the equation

$$(39)\qquad -\frac{(m-2)^3}{m-1}a_0^{m+1}b_0z^2+32S_{m+1}xy=0,$$

that is, a conic determined by the neighborhood of order $2m+1$ of the curve at its $S^m_{1,m-3}$. The locus of the intersection of the lines (37) and (38) is given by the equations

$$(40)\qquad \begin{gathered} mx\rho^{m-1}+(-1)^m\frac{a_0^m}{b_0}y=0,\\ 2[2yS_{m+1}\rho^2+(m-1)(m-2)b_0a_0^{m+1}x]+(m-2)^2a_0^{m+1}b_0\rho z=0. \end{gathered}$$

7. If non-homogeneous coördinates are used, we have

$$(41)\qquad a_0=1,\qquad c_i=0,\qquad a_i=0\qquad (i=1,2,\cdots),$$

$$(42)\qquad y=b_0x^m+b_1x^{m+1}+\cdots+b_\nu x^{m+\nu}+\cdots.$$

Since $OO_{m+1}O_{2m}$ is taken for the triangle of reference, we have

$$b_1=0,\qquad \frac{1}{b_0}U_{10,k}-A_{m,k}=0\quad (k=2,3,\cdots,m).$$

From (41) it follows that

$$b_j=0\qquad\qquad (j=1,\cdots,m).$$

If a point of intersection of (39) and (40) is taken for the unit point, then

$$b_0=\frac{(m-1)^{m-1}}{4^{m-1}m(m-2)^{m-1}},\qquad b_{m+1}=\frac{-(m-1)^{2m-3}}{2^{2m+1}m^2(m-2)^{2m-5}},$$

provided the conic (39) is proper. Hence the canonical expansion of a curve in this case ($m>4$) takes the form [3]

$$(43)\qquad y=\frac{(m-1)^{m-1}}{4^{m-1}m(m-2)^{m-2}}x^m-\frac{(m-1)^{2m-3}}{2^{2m+1}m^2(m-2)^{2m-5}}x^{2m+1}+(2m+2).$$

8. On the contrary, if the conic (39) is degenerate, the canonical expansion for the curve is quite different from (43). In this case C_m has at $S^m_{1,m-3}$ a contact of order $2m+2$ with C if and only if (18), (27), (33),

$$(44)\qquad (m-2)z_1\omega_2^{m+1}+(-1)^m(m-1)\frac{a_0^m}{b_0}x_1y_1y_2\omega_1^m=0,$$

and

$$
\begin{aligned}
&(-1)^m\left(A_{m,m+2} - \frac{a_0^m}{b_0} U_{10,m+2}\right)\left(\frac{y_1y_2}{\omega_3}\right)^2 \\
(45)\quad &+ \begin{bmatrix} m-2 \\ 2 \end{bmatrix}(m-1)(-1)^m b_0 a_0^{m+2}\left\{z_1y_1\left[\frac{y_2}{\omega_3} + \frac{m-2}{m-1}\frac{b_0}{a_0^m}\frac{1}{y_1}\left(\frac{\omega_3}{\omega_1}\right)^{m-2}\right]^2\right. \\
&\qquad\qquad \left. + \left(\frac{y_1y_2}{\omega_3}\right)^2\frac{\omega_1^2\omega_2}{\omega_3^2}\right\} = 0
\end{aligned}
$$

are satisfied. Elimination of y_2 from these equations gives

$$
(46)\qquad \rho^4 = -\left(\frac{m-1}{m-2}\right)^{m-3}\frac{a_0^m yz^{m-5}}{(m-1)b_0x^{m-4}},
$$

$$
(47)\qquad 2\,\frac{m-1}{m-2}\,a_0^m xy + (-1)^m b_0 z\left(m\rho^m x + (-1)^m\frac{a_0^m}{b_0}\rho y\right) = 0,
$$

$$
\begin{aligned}
(48)\quad &\left(\frac{1}{a_0b_0}\right)^2\left(A_{m,m+2} - \frac{a_0^m}{b_0}U_{10,m+2}\right)y - 3\begin{bmatrix} m-1 \\ 3 \end{bmatrix}\frac{a_0^m}{b_0}\,z\left(1 + \frac{(-1)^{m+1}x}{\rho^2 z}\right)^2 \\
&\qquad + 3\begin{bmatrix} m-1 \\ 3 \end{bmatrix}\frac{a_0^m}{b_0}\frac{1}{\rho^2}\left(z + \frac{x}{\rho}\right) = 0.
\end{aligned}
$$

Eliminating ρ from two of (46), (47) and (48), we obtain two curves; each of their common points, except those on $y = 0$, can be taken for an $(m-2)$-ple point of a C_m which has at $S^m_{1,m-3}$ a contact of order $2m + 2$ with C. From (46) and (48) we obtain

$$
\begin{aligned}
&\frac{1}{z^4(z + 2(-1)^m x)^4}\left[B(Ryz - z^2)\frac{yz^{m-5}}{x^{m-4}} - x^2\right]^4 + B^2\cdot\frac{y^2z^{2(m-4)}}{x^{2(m-4)}} \\
(49)\quad &-B\,\frac{yz^{m-5}}{x^{m-4}}\left\{\left[\left(\frac{x}{z + 2(-1)^m x}\right)^2 - 2\,\frac{(Ryz - z^2)Byz^{m-5} - x^{m-2}}{(z^2 + 2(-1)^m xz)x^{m-4}}\right]^2\right. \\
&\qquad \left. -2\left[\frac{(Ryz - z^2)Byz^{m-5} - x^{m-2}}{(z^2 + (-1)^m 2xz)x^{m-4}}\right]^2\right\} = 0,
\end{aligned}
$$

where

$$
B = -\left(\frac{m-1}{m-2}\right)^{m-3}\frac{a_0^m}{(m-1)b_0},
$$

$$
R = \frac{2}{(m-1)(m-2)(m-3)a_0^{m+2}b_0}\left(A_{m,m+2} - \frac{a_0^m}{b_0}U_{10,m+2}\right).
$$

According as $m \equiv r \pmod 4$, $r = 0, 1, 2, 3$, there are four possible cases in eliminating ρ from (46) and (47). In fact, when $m = 4k, 4k + 1, 4k + 2$ or $4k + 3$, we have

$$(50.1)\qquad \left(\frac{1}{Az}\right)^4\left\{2\cdot\frac{m-1}{m-2}Ax + B^k m\,\frac{y^{k-1}z^{k(m-5)+1}}{x^{k(m-4)-1}}\right\}^4 = B\,\frac{yz^{m-5}}{x^{m-4}},$$

$$(50.2)\qquad \left\{\frac{2(m-1)Ax^{k(m-4)}}{(m-2)z(mB^k y^{k-1}z^{k(m-5)} + Ax^{k(m-4)-1})}\right\}^4 = B\,\frac{yz^{m-5}}{x^{m-4}},$$

$$\begin{aligned}(50.3)\qquad &\left[\left(\frac{m-1}{m-2}\right)\frac{2Ax^{k(m-4)}}{my^{k-1}z^{k(m-5)+1}}\right]^4 + \frac{B^2y^2z^{2(m-5)}}{x^{2(m-4)}}\\ &-B\,\frac{yz^{m-5}}{x^{m-4}}\left\{\left[\left(\frac{Ax^{k(m-4)-1}}{mB^k y^{k-1}z^{k(m-5)}}\right)^2 - \left(\frac{m-1}{m-2}\right)\cdot\frac{4Ax^{k(m-4)}}{mB^k y^{k-1}z^{1+k(m-5)}}\right]^2\right.\\ &\left.\qquad - 2\left[\left(\frac{m-1}{m-2}\right)\cdot\frac{2Ax^{k(m-4)}}{mB^k y^{k-1}z^{1+k(m-5)}}\right]\right\} = 0\end{aligned}$$

or

$$\begin{aligned}(50.4)\qquad &\frac{1}{yz^{m-5}}\left(mxB^k\cdot\frac{y^{k-1}z^{k(m-5)}}{Ax^{k(m-4)}}\right)^4 + \frac{B^2yz^{m-5}}{x^{2(m-4)}}\\ &- B\cdot\frac{1}{x^{m-4}}\left\{\left[\left(\frac{m-1}{m-2}\cdot\frac{2x}{z}\right)^2 - 2\cdot\frac{mB^k xy^{k-1}z^{k(m-5)}}{x^{k(m-4)}}\right]^2\right.\\ &\left.\qquad - 2\left[\frac{mB^k xy^{k-1}z^{k(m-5)}}{x^{k(m-4)}}\right]^2\right\} = 0,\end{aligned}$$

where $A = (-1)^m a_0^m/b_0$. These curves are of order $m(m-4)$ with an $[m(m-5)+4]$-ple point at O_{2m} and an $(m-4)$-ple point at O. The curve (49) is of order $4(m-2)$ which has a $4(m-4)$-ple point at O_{2m} with coincident tangents $z = 0$ and a quadruple point at O with coincident tangents $y = 0$ too. The curve (49) intersects the curve (50.i) ($i = 1, 2, 3, 4$), in general, at $4(m-4)(3m-5)+1$ points outside $y = 0$, and each of them can be taken for an $(m-2)$-ple point of one curve of $4(m-4)(3m-4)+1$ C_m's which have at $S_{1,m-3}^m$ a contact of order $2m+2$ with C. Therefore, the behavior of the singularity under consideration is represented by the curve (49). The line l_{2m-1} meets the curve (49) at O, O_{2m} and another point O_{2m+2}:

$$(51)\qquad \left(0,\ (m-1)(m-2)(m-3)a_0^{m+2}b_0,\ 2\left(A_{m,m+2} - \frac{a_0^m}{b_0}U_{10,m+2}\right)\right),$$

a point determined by the neighborhood of order $2m+2$ of the curve at $S_{1,m+3}^m$, provided the conic (39) is degenerate.

9. Taking a point of intersection of the line $O_{m+1}O_{2m+2}$ and the curve (49) for the unit point, we obtain the canonical expansion of the curve in the case in which the conic (39) degenerates but O_{2m+2} does not coincide with O_{2m}:

$$y = b_0 x^m - 3\left[\frac{m-1}{3}\right] b_0^2 x^{2m+2} + (2m+3), \tag{52}$$

where

$$\left(\frac{1}{1+2(-1)^m}\right)^4 b_0^2 + \left(\frac{m-1}{m-2}\right)^{m-3} \cdot \frac{b_0}{m-1}\left\{\left[\left(\frac{1}{1+2(-1)^m}\right)^2 + \frac{2}{1+2(-1)^m}\right]^2 - \frac{2}{(1+2(-1)^m)^2}\right\} + \frac{1}{(m-1)^2}\left(\frac{m-1}{m-2}\right)^{2(m-3)} = 0.$$

According as the conic (39) is proper or not, the $S^m_{1,m-3}$ is said to be of the first or the second species. It may be of interest to investigate the other cases of the singularity $S^m_{1,(m-3)}$ of a curve.

Bibliography

1. E. Bompiani, *Per lo studio proiettivo-differenziale delle singolarità*, Boll. dell'Unione Mat. Ital., vol. 5(1926), pp. 118–120.
2. S. C. Chang, *On the point of inflexion of a plane curve*, this Journal, vol. 9, pp. 823–832.
3. S. C. Chang, *On the singularity S^4_1 of a plane curve.*
4. B. Su, *An extension of Bompiani's osculants for a plane curve with a singular point*, Tôhoku Mathematical Journal, vol. 45(1939), pp. 239–244.
5. B. Su, *Projective differential geometry of singularities of plane curves*, Journal of the Chinese Mathematical Society, vol. 2(1940), pp. 139–151.

National University of Chekiang, Meitan, Kweichow, China.

THE GROWTH OF SOLUTIONS OF A DIFFERENTIAL EQUATION

By Mary L. Boas, R. P. Boas, Jr. and Norman Levinson

It has been shown by D. Caligo [1] that if $A(x)$ is continuous in $0 \leq x < \infty$ and $A(x) = O(x^{-2-\epsilon})$ as $x \to \infty$, with $\epsilon > 0$, then for any solution $y(x)$ of $y''(x) + A(x)y(x) = 0$, $\lim_{x\to\infty} y'(x)$ exists. It follows at once (e.g., by l'Hospital's rule) that $y(x)/x$ has the same limit.

Here we shall give a short proof of a somewhat more general result (Theorem 1), and then prove a still more extensive result (Theorem 2).

1. Theorem 1. *If $A(x)$ and $B(x)$ are continuous in $0 \leq x < \infty$, if*

$$\int_0^\infty x \mid A(x) \mid dx < \infty, \tag{1.0}$$

and if

$$\int_0^{\to\infty} B(x)\, dx \tag{1.1}$$

exists, then for any solution $y(x)$ of

$$y''(x) + A(x)y(x) = B(x), \tag{1.2}$$

$$\lim_{x\to\infty} y'(x) \text{ exists.} \tag{1.3}$$

In the proofs of Theorems 1 and 2 we require the following lemma.

Lemma. *If $f(x)$ is continuous in $0 \leq x < \infty$, if $M(x)$ denotes the maximum of $\mid f(t) \mid$ in $0 \leq t \leq x$, and if for some positive numbers α and x_0*

$$\mid f(x) \mid \leq \alpha + \tfrac{1}{2}M(x) \qquad (x \geq x_0), \tag{1.4}$$

then $f(x)$ is bounded in $0 \leq x < \infty$.

For the proof, we suppose $f(x)$ unbounded, and let λ be a number larger than both 2α and $M(x_0)$. Let x_1 be the greatest lower bound of numbers x such that $\mid f(x) \mid \geq \lambda$; such numbers exist because $f(x)$ is unbounded. Since $f(x)$ is continuous, $f(x_1) = \lambda$; since $\mid f(x) \mid < \lambda$ for $x < x_1$, $M(x_1) = \lambda$; since $M(x_0) < \lambda$, $x_1 > x_0$. From (1.4) with $x = x_1$ we now have $\mid f(x) \mid \leq \alpha + \frac{1}{2}\lambda < \lambda$, which is a contradiction.

Received April 23, 1942; in revised form July 24, 1942.

We now prove Theorem 1. We choose, as we may because of (1.0), a positive x_0 such that

$$\int_{x_0}^{\infty} x \mid A(x) \mid dx < \tfrac{1}{2}. \tag{1.5}$$

Clearly

$$y(x) = y(x_0) + \int_{x_0}^{x} y'(t)\, dt. \tag{1.6}$$

From (1.2), we have

$$y''(x) = -A(x)y(x) + B(x).$$

Substituting for $y(x)$ from (1.6), we obtain

$$y''(x) = -y(x_0)A(x) - A(x) \int_{x_0}^{x} y'(t)\, dt + B(x);$$

and, on integrating this relation,

$$\begin{aligned} y'(x) = y'(x_0) - y(x_0) \int_{x_0}^{x} A(t)\, dt \\ - \int_{x_0}^{x} A(u)\, du \int_{x_0}^{u} y'(t)\, dt + \int_{x_0}^{x} B(t)\, dt. \end{aligned} \tag{1.7}$$

We define

$$\alpha = \mid y'(x_0) \mid + \mid y(x_0) \mid \int_{x_0}^{\infty} \mid A(t) \mid dt + \max_{x_0 \leq x < \infty} \left| \int_{x_0}^{x} B(t)\, dt \right|,$$

$$M(x) = \max_{0 \leq t \leq x} \mid y'(t) \mid.$$

From (1.7) we now obtain

$$\begin{aligned} \mid y'(x) \mid &\leq \alpha + M(x) \int_{x_0}^{x} \mid A(u) \mid du \int_{x_0}^{u} dt \\ &\leq \alpha + M(x) \int_{x_0}^{x} u \mid A(u) \mid du. \end{aligned}$$

Using (1.5), we thus have

$$\mid y'(x) \mid \leq \alpha + \tfrac{1}{2} M(x).$$

From the lemma we conclude that $y'(x)$ is bounded in $0 \leq x < \infty$. But if $y'(x)$ is bounded, it follows from (1.0) that the integrals involving $A(x)$ on the right of (1.7) approach limits as $x \to \infty$; the integral involving $B(x)$ approaches a limit by (1.1); consequently $\lim_{x \to \infty} y'(x)$ exists. This completes the proof of Theorem 1.

2. Here we consider the differential equation (1.2) under more general conditions. Let

$$A_1(x) = \frac{|A(x)| + A(x)}{2},$$

$$A_2(x) = \frac{|A(x)| - A(x)}{2}.$$

Then clearly $A_1(x)$ and $A_2(x)$ are non-negative and $A(x) = A_1(x) - A_2(x)$.

THEOREM 2. *Suppose that $A(x)$ and $B(x)$ are continuous, that*

$$\limsup_{x\to\infty} x \int_x^\infty A_1(t)\,dt < 1, \tag{2.0}$$

$$\int_0^\infty x A_2(x)\,dx < \infty, \tag{2.1}$$

and

$$\int_0^{\to\infty} B(x)\,dx \qquad \textit{exists.} \tag{2.2}$$

(2.0) will, in particular, be true if $\limsup_{x\to\infty} x^2 A_1(x) < 1$. Then if $y''(x) + A(x)y(x) = B(x)$,

$$\lim_{x\to\infty} y'(x) \tag{2.3}$$

exists. Moreover, if

$$\int_0^\infty x A_1(x)\,dx = \infty, \tag{2.4}$$

then

$$\lim_{x\to\infty} y'(x) = 0. \tag{2.5}$$

3. **Proof of** (2.3) **when** $y'(x)$ **does not change sign for large** x. Without loss of generality we assume that $y'(x) \geq 0$ for $x \geq x_0$ (since replacing $y(x)$ by $-y(x)$ does not affect our hypotheses). Suppose that x_0 is so large that

$$\int_{x_0}^\infty x A_2(x)\,dx < \tfrac{1}{2}. \tag{3.0}$$

Using (1.7), we have

$$y'(x) = y'(x_0) - y(x_0) \int_{x_0}^{x} A(t)\,dt - \int_{x_0}^{x} A(u)\,du \int_{x_0}^{u} y'(t)\,dt + \int_{x_0}^{x} B(t)\,dt. \tag{3.1}$$

Since $y'(t) \geq 0$ for $t \geq x_0$, and $-A(t) = A_2(t) - A_1(t) \leq A_2(t)$, we then have for $x \geq x_0$

$$| y'(x) | \leq | y'(x_0) | + | y(x_0) | \int_{x_0}^{\infty} | A(t) |\,dt + \int_{x_0}^{x} A_2(u)\,du \int_{x_0}^{u} y'(t)\,dt + B, \tag{3.2}$$

where

$$B = \max_{x_0 \leq x < \infty} \left| \int_{x_0}^{x} B(u)\,du \right|;$$

the integral $\int^{\infty} | A(t) |\,dt$ converges by (2.0) and (2.1). We define

$$\alpha = | y'(x_0) | + | y(x_0) | \int_{x_0}^{\infty} | A(t) |\,dt + B,$$

$$M(x) = \max_{0 \leq t \leq x} | y'(t) |.$$

Then, from (3.2) and (3.0),

$$| y'(x) | \leq \alpha + M(x) \int_{x_0}^{x} A_2(u)\,du \int_{x_0}^{u} dt$$

$$\leq \alpha + M(x) \int_{x_0}^{x} uA_2(u)\,du$$

$$\leq \alpha + \tfrac{1}{2}M(x).$$

By the lemma, $y'(x)$ is bounded in $0 \leq x < \infty$. In (3.1), write $A(u) = A_1(u) - A_2(u)$. Then, since $y'(x)$ is bounded, all the terms on the right of (3.1) approach limits as $x \to \infty$, with the possible exception of

$$- \int_{x_0}^{x} A_1(u)\,du \int_{x_0}^{u} y'(t)\,dt.$$

But, since $y'(t) \geq 0$ for $t \geq x_0$, and $A_1(u) \geq 0$, this term is a non-increasing function of x; it is bounded below because all the other terms in (3.1) are bounded. Hence it also approaches a limit.

4. **Proof of (2.3) when $y'(x)$ changes sign an infinite number of times as** $x \to \infty$. We consider two cases.

Case 1. Here we prove (2.3) when $y'(x)$ is bounded. If $\lim\limits_{x\to\infty} y'(x)$ does not exist, we can assume without loss of generality that

$$\limsup_{x\to\infty} | y'(x) | = \limsup_{x\to\infty} y'(x) = M > 0. \tag{4.0}$$

Choose a sequence of points x_n, $x_n \to \infty$, such that $y'(x_n) \to M$. By (2.0) there is a number c, $0 < c < 1$, such that

$$x \int_x^\infty A_1(t)\, dt < c \tag{4.1}$$

for large x. Let a_n denote the first point to the left of x_n at which $y'(x) = 0$. We may suppose that (4.1) holds for $x \geq a_1$.

We now write (3.1) with a_n in place of x_0 and x_n in place of x. Since $y'(a_n) = 0$, we obtain

$$y'(x_n) = -y(a_n) \int_{a_n}^{x_n} A(u)\, du - \int_{a_n}^{x_n} A(u)\, du \int_{a_n}^{u} y'(t)\, dt + \int_{a_n}^{x_n} B(u)\, du.$$

Since $y'(t) \geq 0$ in $a_n \leq t \leq x_n$, it follows that

$$\begin{aligned} y'(x_n) \leq | y(a_n) | \int_{a_n}^{x_n} A_1(u)\, du + | y(a_n) | \int_{a_n}^{x_n} A_2(u)\, du \\ + \int_{a_n}^{x_n} A_2(u)\, du \int_{a_n}^{u} y'(t)\, dt + \left| \int_{a_n}^{x_n} B(u)\, du \right|. \end{aligned} \tag{4.2}$$

Of the terms on the right of (4.2) the last approaches zero by (2.2); the third approaches zero by (2.1), since $y'(t)$ is bounded. For the second, we have

$$\begin{aligned} \limsup_{n\to\infty} | y(a_n) | \int_{a_n}^{x_n} A_2(u)\, du &\leq \limsup_{n\to\infty} \frac{| y(a_n) |}{a_n} \int_{a_n}^{x_n} u A_2(u)\, du \\ &\leq \limsup_{x\to\infty} | y'(x) | \int_x^\infty u A_2(u)\, du = 0. \end{aligned}$$

For the first, we have, by (4.1),

$$\begin{aligned} \limsup_{n\to\infty} | y(a_n) | \int_{a_n}^{x_n} A_1(u)\, du &\leq \limsup_{n\to\infty} | y(a_n) | \int_{a_n}^{\infty} A_1(u)\, du \\ &\leq c \limsup_{n\to\infty} \frac{| y(a_n) |}{a_n} \\ &\leq c \limsup_{x\to\infty} | y'(x) | = cM. \end{aligned}$$

Consequently, (4.2) implies

$$\limsup_{n\to\infty} y'(x_n) \le cM < M,$$

contradicting the choice of the points x_n.

Case 2. Here we show that $y'(x)$ cannot be unbounded. Assuming that $y'(x)$ is unbounded, we can choose a sequence of points x_n, $x_n \to \infty$, such that

$$|y'(x_n)| \ge |y'(x)|, \qquad 0 < x \le x_n; \tag{4.3}$$

$y'(x_n)$ has the same sign (which we may assume to be positive) for all n; and $y'(x_n) \to \infty$. Let a_n be the first point to the left of x_n at which $y'(x) = 0$. We may suppose that (4.1) holds for $x \ge a_1$. Now consider (4.2) again. By (4.3),

$$|y(a_n)| \le \int_0^{a_n} |y'(t)|\,dt + |y(0)| \le a_n |y'(x_n)| + |y(0)|. \tag{4.4}$$

We use (4.1) and (4.4) in the first two terms on the right of (4.2), and (4.3) in the third. We obtain

$$y'(x_n) \le y'(x_n)\left\{a_n \int_{a_n}^{x_n} A_1(u)\,du + a_n \int_{a_n}^{x_n} A_2(u)\,du + \int_{a_n}^{x_n} uA_2(u)\,du\right\}$$
$$+ |y(0)| \int_{a_n}^{x_n} [A_1(u) + A_2(u)]\,du + \left|\int_{a_n}^{x_n} B(u)\,du\right|.$$

Using (4.1), we then have

$$\begin{aligned} y'(x_n) &\le y'(x_n)\left\{c + 2\int_{a_n}^{\infty} uA_2(u)\,du\right\} \\ &\quad + |y(0)| \int_{a_n}^{\infty} [A_1(u) + A_2(u)]\,du + \left|\int_{a_n}^{x_n} B(u)\,du\right|. \end{aligned} \tag{4.5}$$

Since all the integrals on the right of (4.5) approach zero as $n \to \infty$, and $0 < c < 1$, we reach a contradiction. Consequently, $y'(x)$ cannot be unbounded.

5. We now prove (2.5) when (2.4) holds. Suppose that (2.4) is true, but that $y'(x)$ does not approach zero. Without loss of generality we suppose that

$$\lim_{x\to\infty} y'(x) = 2a > 0. \tag{5.0}$$

We now refer to (3.1) and let $x \to \infty$. Since $y'(x)$ is bounded, all the terms on the right of (3.1) approach limits, with the possible exception of

$$-\int_{x_0}^{x} A_1(u)\,du \int_{x_0}^{u} y'(t)\,dt. \tag{5.1}$$

Since $y'(x)$ approaches a limit, this term must approach a limit also. But, by (5.0), for large x_0 and $x > x_0$ we have $y'(x) > a$,

$$\int_{x_0}^{u} y'(t)\, dt > a(u - x_0),$$

$$\int_{x_0}^{x} A_1(u)\, du \int_{x_0}^{u} y'(t)\, dt > a \int_{x_0}^{x} (u - x_0)A_1(u)\, du;$$

by (2.4), the last integral becomes infinite as $x \to \infty$, so that (5.1) cannot approach a limit. This contradiction establishes (2.5).

Bibliography

1. D. Caligo, *Comportamento asintotico degli integrali dell'equazione* $y''(x) + A(x)y(x) = 0$, *nell'ipotesi* $\lim_{x\to\infty} A(x) = 0$, Bollettino della Unione Matematica Italiana (2), vol. 3(1941), pp. 286–295.

Duke University, U. S. Navy Pre-Flight School, Chapel Hill, N. C., and Massachusetts Institute of Technology.

THE INFLUENCE OF GAPS ON DENSITY OF INTEGERS

By R. Salem and D. C. Spencer

1. It is a well-known fact in the theory of sets of points that if a set E is such that the part of E contained in every interval (a, b) has a measure less than $\theta(b - a)$ $(0 < \theta < 1)$, then E is of measure zero.

The analogous theorem obtained by replacing in the preceding statement "set of points" by "sequence of integers" and "measure" by "number of integers" is false (the length $b - a$ being, for obvious reasons, not less than a certain lower bound l_0), as is seen by the trivial example of the sequence of even integers. It is remarkable, however, that the picture changes completely if the missing integers in every interval are supposed to be consecutive—that is to say, if they form a gap. It has been found interesting to investigate this subject.

We consider infinite sequences of increasing non-negative integers $u_1, u_2, \cdots, u_n, \cdots$, and we suppose always that the first number of the sequence $u_1 = 0$. By $\nu(m)$ we denote the number of terms of the sequence which are not greater than m. If $\underline{\lim}\, \nu(m)/m > 0$, we say that the sequence has a positive density. If $\nu(m)/m$ tends to zero, we say that the density of the sequence is zero; if, in this case, $\nu(m)/m = O(\varphi(m))$, we say that the density is at most of order $\varphi(m)$; if $\varphi(m) = O(\nu(m)/m)$, we say that the density is at least of order $\varphi(m)$; and if the last two relations hold simultaneously, that the density is of order $\varphi(m)$.

Let $\omega(x)$ be a non-negative, non-decreasing function of the non-negative variable x. We say that a sequence has the complete gap property with respect to ω if in any closed interval $(a, a + l)$ $(a \geq 0, l \geq l_0 > 0)$ there exists an open interval of length not less than $\omega(l)$ which contains no point of the sequence. The open interval will be called a gap. That there exist sequences having a complete gap property with respect to some function $\omega(x)$ is not difficult to see by considering simple sequences like $u_n = (n - 1)^2$.

It should be observed that by imposing on the intervals $(a, a - l)$ certain restrictions which may seem trivial prima facie, such as "the end points of the intervals must belong to the sequence", the character of the problem may change substantially and the results may be quite different. We do not propose to investigate these alternatives here.

The following theorems will be proved:

Theorem I. *If the integral*

$$\int_1^\infty \frac{\omega(x)}{x^2}\, dx$$

Received May 6, 1942.

is divergent, any sequence having the complete gap property with respect to $\omega(x)$ has density zero.

THEOREM II. *Given any increasing function $\omega(x)$ such that $\omega(x)/x^2$ decreases and the integral*

$$\int_1^\infty \frac{\omega(x)}{x^2}\,dx$$

converges, there exists a sequence of positive density having the complete gap property with respect to $\omega(x)$.

THEOREM III. *If $\omega(x) = \theta x$, where $0 < \theta < \frac{1}{3}$, any sequence having the complete gap property with respect to $\omega(x)$ has a density of order $m^{-\alpha}$ at most, where*

$$\alpha = \log\left(\frac{1-2\theta}{1-3\theta}\right) \Big/ \log 2\left(\frac{1-2\theta}{1-3\theta}\right).$$

THEOREM IV. *If $\omega(x) = \theta x$, where $0 < \theta < \frac{1}{3}$, there exists a sequence having the complete gap property with respect to $\omega(x)$ and having a density of order $m^{-\alpha}$, where α has the value stated in Theorem* III.

THEOREM V. *If $\omega(x) = \theta x$, where $\frac{1}{3} \le \theta < \frac{1}{2}$, any sequence having the complete gap property with respect to $\omega(x)$ has a density at most of order* $\log m/m$.

THEOREM VI. *If $\omega(x) = \theta x$, where $\frac{1}{3} \le \theta < \frac{1}{2}$, there exists a sequence having the complete gap property with respect to $\omega(x)$ and having a density of order* $\log m/m$.

THEOREM VII. *If $\omega(x) \ge \frac{1}{2}x$, there is no infinite sequence having the complete gap property with respect to $\omega(x)$.*

We shall not attempt to give here a systematic treatment of the intermediate cases $\omega(x) = \epsilon(x)$, $\omega(x) = \frac{1}{3}x - \epsilon(x)$, $\omega(x) = \frac{1}{2}x - \epsilon(x)$, $\epsilon(x)$ being positive and of order $o(x)$ as $x \to \infty$. It would be interesting to investigate these cases in detail, and to determine to what extent the methods of this paper are effective in treating them.

2. **Proof of Theorem** I. Let u_1 $(=0)$, u_2, $\cdots$, u_n, $\cdots$ be a sequence having the complete gap property with respect to ω. In the closed interval $(0, m)$ there is, if $m \ge l_0$, a gap $G_1 \ge \omega(m)$. Denoting the subsequence contained in $(0, m)$ by U_m, we see that U_m is entirely included in two closed intervals. One of them has a length not less than $\frac{1}{2}(m - G_1)$ and so, if $\frac{1}{2}(m - G_1) \ge l_0$, contains a gap $G_2 \ge \omega(\frac{1}{2}m - \frac{1}{2}G_1)$. U_m is thus entirely included in three closed intervals, one of which has a length not less than $\frac{1}{3}(m - G_1 - G_2)$. If this number is not less

than l_0, we can continue. After k operations we obtain a sequence of gaps $G_1, G_2, \cdots, G_k$, such that

$$G_i \geq \omega\left(\frac{m - G_1 - G_2 - \cdots - G_{i-1}}{i}\right) \qquad (i = 1, 2, \cdots, k)$$

provided

$$\frac{m - G_1 - G_2 - \cdots - G_{k-1}}{k} \geq l_0 . \tag{2.1}$$

But each open interval of length $G_i \geq 1$ contains $G_i - 1$ lattice points representing integers, and so

$$\nu(m) \leq [m] + 1 - \sum_{i=1}^{k} (G_i - 1) \leq m + k + 1 - \sum_{i=1}^{k} G_i .$$

That is to say

$$\frac{m - G_1 - G_2 - \cdots - G_k}{k} \geq \frac{\nu(m)}{k} - \frac{k+1}{k},$$

and hence if we choose k such that $\nu(m)/k - 2 \geq l_0$, the inequality (2.1) will be satisfied. We take

$$k = \left[\frac{\nu(m)}{l_0 + 2}\right],$$

$[z]$ denoting as usual the integral part of z. Then (for m large enough)

$$m - G_1 - G_2 - \cdots - G_k \geq \nu(m) - \frac{\nu(m)}{l_0 + 2} - 1 \geq \frac{\nu(m)}{2}. \tag{2.2}$$

From this inequality we have

$$G_1 \geq \omega(m) \geq \omega\left(\frac{\nu(m)}{2}\right),$$

$$G_2 \geq \omega\left(\frac{m - G_1}{2}\right) \geq \omega\left(\frac{\nu(m)}{2\cdot 2}\right),$$

$$\cdots \qquad \cdots \qquad \cdots$$

$$G_k \geq \omega\left(\frac{m - G_1 - \cdots - G_{k-1}}{k}\right) \geq \omega\left(\frac{\nu(m)}{2k}\right).$$

This gives immediately, $\omega(x)$ being non-decreasing,

$$G_1 + G_2 + \cdots + G_k \geq \int_1^{k+1} \omega\left(\frac{\nu(m)}{2y}\right) dy \geq \int_1^{\frac{\nu(m)}{l_0+2}} \omega\left(\frac{\nu(m)}{2y}\right) dy.$$

Let $y = \nu(m)/2x$; then

$$G_1 + \cdots + G_k \geq \frac{\nu(m)}{2} \int_{\frac{1}{2}(l_0+2)}^{\frac{1}{2}\nu(m)} \frac{\omega(x)}{x^2}\, dx,$$

and so by (2.2)

$$m \geq \frac{\nu(m)}{2}\left[1 + \int_{\frac{1}{2}(l_0+2)}^{\frac{1}{2}\nu(m)} \frac{\omega(x)}{x^2}\, dx\right].$$

That is to say

$$\frac{\nu(m)}{m} \leq \frac{2}{1 + \int_{\frac{1}{2}(l_0+2)}^{\frac{1}{2}\nu(m)} \frac{\omega(x)}{x^2}\, dx},$$

a result which implies Theorem I.

Theorem III will show that this inequality, sufficient to prove Theorem I, does not give any precise information as to the order of the density when the density is zero.

Before proving Theorem II we shall investigate some of the properties of a particular type of sequence of integers.

3. **Study of a particular type of sequence of integers.** We define a sequence U:

$$u(1) = 0, \qquad u(2), \qquad u(3), \qquad \cdots, \qquad u(n), \qquad \cdots$$

by the series

$$u(n+1) = ng_0 + \left[\frac{n}{2}\right]g_1 + \left[\frac{n}{2^2}\right]g_2 + \cdots + \left[\frac{n}{2^p}\right]g_p + \cdots,$$

where the g_p are integers satisfying

$$g_0 \geq 1; \qquad g_p \geq 0 \qquad (p \geq 1),$$

and where $[z]$ denotes the integral part of z. The series obviously breaks off and so contains only a finite number of terms. Since $g_0 \geq 1$, $u(n)$ is plainly a strictly increasing function of n.

If the series $\sum_0^\infty \frac{g_p}{2^p}$ converges, let σ be its sum; we have $u(n+1) \leq n\sigma$ and so the sequence is of positive density. If this series diverges, it is easily seen that $u(n)/n$ increases infinitely with n, and so the sequence has density zero.

If n is written in the dyadic system

$$n = \epsilon_0 2^m + \epsilon_1 2^{m-1} + \cdots + \epsilon_m \qquad (\epsilon_i = 0 \text{ or } 1),$$

we plainly have

$$(3.1)\qquad u(n+1) = \epsilon_0 u(2^m+1) + \epsilon_1 u(2^{m-1}+1) + \cdots + \epsilon_{m-1}u(3) + \epsilon_m u(2).$$

That is to say, $u(n+1)$ is derived from n by the replacement, in the dyadic development of n, of 2^r by $u(2^r+1)$.

An integer m being fixed, let us now consider the set S_m of the first 2^m numbers of the sequence: $u(1)$, $u(2)$, $\cdots$, $u(2^m)$. We shall prove that the sequence U consists of an infinity of sets S_m , S_m^1 , S_m^2 , $\cdots$ each of them being the result of a mere translation of S_m , and two successive sets being separated by a gap. In fact, every natural number can be written uniquely in the form $q2^m + k$, q and k being integers, and k belonging to the sequence 1, 2, $\cdots$, 2^m. Hence

$$u(q2^m + k) = u(k) + T_q ,$$

where T_q is equal to $u(q2^m + 1)$ and so depends on q but not on k. Thus S_m^q is the translation of S_m by T_q . The sets S_m^q will be called sets of the type S_m .

We observe that for every fixed m the sequence U can be split into a family of sets of type S_m , of 2^m terms each, that S_{m+1} consists of $S_m + S_m^1$, and that between S_m and S_m^1 there is a gap of length

$$\begin{aligned} u(2^m+1) - u(2^m) &= 2^m g_0 + 2^{m-1}g_1 + \cdots + g_m \\ &\qquad - [(2^m-1)g_0 + (2^{m-1}-1)g_1 + \cdots + (2-1)g_{m-1}] \\ &= g_0 + g_1 + \cdots + g_{m-1} + g_m = \gamma_{m+1} , \end{aligned}$$

say. Furthermore between any two sets of the type S_m there is a gap at least equal to γ_{m+1} , for it is easily seen that

$$u(q2^m + 1) - u(q2^m) \geq g_0 + g_1 + \cdots + g_m .$$

We have proved incidentally that every set S_m^q has a center of symmetry and that there is a gap of length γ_m which lies in the middle of the set. The distance between the first and last terms of the set S_m^q is of course $u(2^m)$.

We proceed now to study the complete gap property of the sequence U by proving two lemmas. Let an interval of length l be given and let us suppose that

$$(3.2)\qquad u(2^m) + 2\gamma_m \leq l < u(2^{m+1}) + 2\gamma_{m+1} .$$

We can then write

$$l = u(2^m) + 2\gamma_m + x,$$

where

$$0 \leq x < 2\gamma_{m+1} - 2\gamma_m + u(2^{m+1}) - u(2^m).$$

We begin by proving the following result:

LEMMA I. *In every interval of length l satisfying* (3.2) *there is a gap which is at least equal to*

$$G(l) = \min (\gamma_m + \tfrac{1}{2}x, \gamma_{m+1}). \tag{3.3}$$

Let $a \geq 0$ and $b = a + l$ be the end points of the given interval. We say that a point lies in a set S_m^q if it lies in the closed interval whose end points are the first and the last points of S_m^q. We consider the following different cases:

(1) *Neither a nor b lies in a set of the type S_m.* (i) If they lie in the same gap separating two sets of the family, there are no points of the sequence in (a, b) and so the gap is at least $2\gamma_m + x > \gamma_m + \frac{1}{2}x \geq G(l)$. (ii) If they lie in two different gaps separating sets of the type S_m and if there are at least two sets of the type S_m between a and b, then the interval (a, b) contains a gap not less than $\gamma_{m+1} \geq G(l)$. (iii) If they lie in two different gaps separating sets of the type S_m and if there is only one set of the type S_m between a and b, let α be the distance between a and the first point of the set, and β the distance between b and the last point of the set. Then $\alpha + \beta = l - u(2^m) = 2\gamma_m + x$ and hence at least one of the gaps α, β is not less than $\gamma_m + \frac{1}{2}x \geq G(l)$.

(2) *Both a and b lie in a set of the type S_m.* The points a and b obviously cannot be in the same set, and thus there exists in (a, b) a gap not less than $\gamma_{m+1} \geq G(l)$.

(3) *One of the points a, b lies in a set of the type S_m, the other in a gap separating two sets of the family.* (i) If this gap is not contiguous to the set, there is in (a, b) a gap not less than $\gamma_{m+1} \geq G(l)$. (ii) If it is contiguous to the set, then the distance between the end point of (a, b) not lying in the set and the nearest point of the set is at least $l - u(2^m) = 2\gamma_m + x > \gamma_m + \frac{1}{2}x \geq G(l)$. This completes the proof.

We add that the consideration of the hypothesis (1) (ii) in the case in which the center of the interval (a, b) is the center of a set of the type S_{m+1} and the consideration of the hypothesis (1) (iii) in the case in which the center of the interval (a, b) is the center of a set of the type S_m show that the result (3.3) is a best possible one when the length of the interval satisfies (3.2).

From this result we deduce the following lemma:

LEMMA II. *Suppose that $\omega(x)$ is a non-negative, differentiable, non-decreasing function for $x \geq 0$, and that $\omega'(x) \leq \frac{1}{2}$ for $x \geq x_0$. Then a necessary and sufficient condition that the sequence U should have the complete gap property with respect to $\omega(x)$ is that for $m \geq m_0$*

$$\gamma_m \geq \omega[u(2^m) + 2\gamma_m]. \tag{3.4}$$

The reason for the restriction $\omega'(x) \leq \frac{1}{2}$ will appear below in the proof of Theorem VII.

The necessity of the condition follows from (3.2), (3.3) and the remark at the

end of the proof of Lemma I. To prove that the condition is sufficient we have only to prove that it implies

$$\gamma_m + \tfrac{1}{2}x \geq \omega[u(2^m) + 2\gamma_m + x]$$

for $0 < x \leq 2\gamma_{m+1} - 2\gamma_m$, and this follows immediately from the hypothesis $\omega'(y) \leq \frac{1}{2}$ if y is large enough.

4. **Proof of Theorem** II. We observe first that if the sequence U has positive density and if

$$\sum_0^\infty \frac{g_p}{2^p} = \sigma,$$

then

$$u(2^m) + 2\gamma_m = (2^m + 1)g_0 + (2^{m-1} + 1)g_1 + \cdots + 3g_{m-1}$$
$$< 2^{m+1}g_0 + \cdots + 2^2 g_{m-1} < 2^{m+1}\sigma.$$

This shows immediately, without use of Lemma II, that if $\gamma_m \geq \omega(2^{m+2}\sigma)$, the sequence will have the complete gap property with respect to $\omega(x)$. For $G(l)$ is at least γ_m for $2^{m+1}\sigma \leq l < 2^{m+2}\sigma$.

Now let $\omega(x)$ be a positive increasing function for $x \geq 0$ such that $\omega(x)/x^2$ decreases and such that

$$\int_1^\infty \frac{\omega(x)}{x^2}\,dx$$

converges. Then the series

$$\sum_1^\infty \frac{\omega(n)}{n^2}, \qquad \sum_0^\infty \frac{\omega(2^m)}{2^m}$$

converge.

We define the sequence U by

$$u(n + 1) = ng_0 + \left[\frac{n}{2}\right]g_1 + \cdots + \left[\frac{n}{2^p}\right]g_p + \cdots,$$

where

$$g_0 = 1,\ g_p = 0 \text{ for } 1 \leq p \leq h - 1,$$

$$g_p = [\omega(2^{p+4})] + 1 \text{ for } p \geq h,$$

h being taken large enough to insure that

$$\sum_h^\infty \frac{[\omega(2^{m+4})] + 1}{2^m} < 1.$$

We then have

$$\sigma = \sum_0^\infty \frac{g_\nu}{2^\nu} < 2$$

and

$$\gamma_m \geq g_{m-1} > \omega(2^{m+3}) > \omega(2^{m+2}\sigma)$$

for $m \geq h + 1$. Hence the sequence has the complete gap property with respect to $\omega(x)$ for $l \geq l_0 = 2^{h+3}$. We note that by taking l_0 large enough we can have a density as near to 1 as we please.

5. **Proof of Theorem** III. Let $u_1(= 0), u_2, \cdots, u_n, \cdots$ be a sequence having the complete gap property with respect to θx for $l \geq l_0$, where $0 < \theta < \frac{1}{3}$.

LEMMA III. *Let u, v be two numbers of the sequence satisfying*

$$v - u \geq l_0, \tag{5.1}$$

$$u \geq \frac{\theta v}{1 - \theta}. \tag{5.2}$$

Then in the interval (u, v) there is a gap of length not less than

$$(v - u)\frac{\theta}{1 - 2\theta}.$$

Suppose not, and let Γ be the length of the maximum gap in (u, v). The inequality $u \geq \theta v/(1 - \theta)$ is equivalent to $u \geq \theta(v - u)/(1 - 2\theta)$, and hence we have $u > \Gamma$. Consider now the interval $(u - \Gamma, v + \Gamma)$. It must contain a gap of length not less than $\theta(v - u + 2\Gamma)$. But this gap cannot be greater than Γ since u, v are points of the sequence, and so $\Gamma \geq \theta(v - u + 2\Gamma)$; that is, $\Gamma \geq \theta(v - u)/(1 - 2\theta)$, contrary to hypothesis, and the lemma is proved.

It will be useful to observe in the applications of the lemma that if $u \geq \theta v/(1 - \theta)$, we have a fortiori $u' \geq \theta v'(1 - \theta)$ if $u \leq u' < v' \leq v$.

We shall henceforth write, for the sake of brevity,

$$\rho = \frac{\theta}{1 - 2\theta},$$

where ρ satisfies $0 < \rho < 1$ since $0 < \theta < \frac{1}{3}$.

LEMMA IV. *Let $G(x)$ be a function equal to zero for $0 \leq x < l_0$ and to ρx for $x \geq l_0$. Then if $a_1, a_2, \cdots, a_k$ are non-negative numbers such that*

$$\frac{a_1 + a_2 + \cdots + a_k}{k} \geq l_0 \tag{5.3}$$

we have

$$G(a_1) + G(a_2) + \cdots + G(a_k) \geq \rho(a_1 + a_2 + \cdots + a_k - \overline{k - 1}\, l_0).$$

It is impossible by (5.3) that all numbers a_i be less than l_0. Suppose that we arrange them in order of magnitude, and that

$$a_1 \leq a_2 \leq \cdots \leq a_h < l_0 \leq a_{h+1} \leq \cdots \leq a_k .$$

Then

$$\sum_{i=1}^{k} G(a_i) = \rho(a_{h+1} + \cdots + a_k) \geq \rho(\sum_{i=1}^{k} a_i - hl_0)$$

$$\geq \rho(\sum_{i=1}^{k} a_i - \overline{k - 1}\, l_0).$$

LEMMA V. *Let u, v be two numbers of the sequence satisfying the inequality* (5.2), *and write $v - u = m$. Let $\mu(m)$ be the number of terms of the sequence in the closed interval (u, v). Then we have $\mu(m) < Cm^\beta$, where $\beta = \log 2/\log [2/(1 - \rho)]$ and $C = C(\theta, l_0)$ is a number depending on θ and l_0 only.*

In fact, let Γ_0 be the gap of maximum length in (u, v), or one such gap; this is tacitly understood hereafter. By Lemma III we have, if $m \geq l_0$ ($G(x)$ being the function defined in Lemma IV), $\Gamma_0 \geq G(m) = \rho m$. Let Γ_1^1 and Γ_2^1 be the lengths of the maximum gaps in the two intervals of lengths m_1^1 and m_2^1, contiguous to Γ_0. We have by Lemma III $\Gamma_1^1 \geq G(m_1^1)$, $\Gamma_2^1 \geq G(m_2^1)$ and by Lemma IV, if $m_1^1 + m_2^1 \geq 2l_0$,

$$\Gamma_1^1 + \Gamma_2^1 \geq \rho(m_1^1 + m_2^1 - l_0) = \rho(m - \Gamma_0 - l_0).$$

We have now four intervals of lengths m_1^2, m_2^2, m_3^2, m_4^2, contiguous to Γ_0, Γ_1^1, Γ_2^1 (some of them may reduce to points), and arguing in the same way we have, if $m_1^2 + m_2^2 + m_3^2 + m_4^2 \geq 2^2 l_0$,

$$\Gamma_1^2 + \Gamma_2^2 + \Gamma_3^2 + \Gamma_4^2 \geq \rho(m - \Gamma_0 - \Gamma_1^1 - \Gamma_2^1 - 3l_0).$$

In general, writing $\Gamma_1^k + \Gamma_2^k + \cdots + \Gamma_{2^k}^k = r_k$, we have

$$r_k \geq \rho(m - r_0 - r_1 - \cdots - r_{k-1} - \overline{2^k - 1}\, l_0) \tag{5.4}$$

provided that

$$m - (r_0 + r_1 + \cdots + r_{k-1}) \geq 2^k l_0 . \tag{5.5}$$

Now each open interval of length $\Gamma \geq 1$ contains $\Gamma - 1$ lattice points representing integers. Hence

$$\mu(m) \leq m + 1 - \sum_{j=0}^{k} (r_j - 2^j) = m + 1 - (r_0 + \cdots + r_k) + 2^{k+1} - 1,$$

that is,

$$(5.6) \qquad m - (r_0 + r_1 + \cdots + r_k) \geq \mu(m) - 2^{k+1},$$

and the inequality (5.5) will thus be satisfied if we take

$$(5.7) \qquad \mu(m) \geq 2^k(l_0 + 2).$$

We fix an integer k satisfying (5.7), and write

$$s_k = r_0 + r_1 + \cdots + r_k .$$

We have then by (5.4)

$$s_k - s_{k-1} \geq \rho m - \rho s_{k-1} - \rho l_0 2^k,$$

or

$$(5.8) \qquad s_k - (1 - \rho)s_{k-1} \geq \rho m - \rho l_0 2^k;$$

and in the same way

$$(5.9) \qquad \begin{aligned} s_{k-1} - (1 - \rho)s_{k-2} &\geq \rho m - \rho l_0 2^{k-1}, \\ \cdots \quad &\cdots \quad \cdots \\ s_1 - (1 - \rho)s_0 &\geq \rho m - \rho l_0 2, \\ s_0 &\geq \rho m . \end{aligned}$$

Multiplying the inequalities (5.9) by $(1 - \rho)$, $(1 - \rho)^2$, $\cdots$, $(1 - \rho)^k$ respectively and adding to (5.8) we obtain

$$s_k \geq \rho m \sum_{i=0}^{k} (1 - \rho)^i - \rho l_0[2^k + 2^{k-1}(1 - \rho) + \cdots + 2(1 - \rho)^{k-1}],$$

and so by (5.6)

$$m - \mu(m) + 2^{k+1} \geq \rho m\left[\frac{1 - (1 - \rho)^{k+1}}{\rho}\right] - \rho l_0 2^k \sum_0^\infty \left(\frac{1 - \rho}{2}\right)^n,$$

that is to say,

$$\mu(m) \leq m(1 - \rho)^{k+1} + (\rho l_0 + 1)2^{k+1}.$$

We now fix k as the greatest integer for which $2^{k+2}(l_0 + 2) \leq \mu(m)$. Then the inequality (5.7) is satisfied a fortiori and we have

$$(\rho l_0 + 1)2^{k+1} < (l_0 + 2)2^{k+1} \leq \tfrac{1}{2}\mu(m).$$

Hence $\mu(m) \leq 2m(1-\rho)^{k+1}$, where (by definition of k) $2^{k+3} > \mu(m)/(l_0 + 2)$.

The last two inequalities give us the order of $\mu(m)$. We have

$$\log \mu(m) \leq \log 2 + \log m - (k+1) \log \frac{1}{1-\rho},$$

$$k + 1 > \frac{\log \mu(m) - \log (l_0 + 2)}{\log 2} - 2,$$

and so

$$\log \mu(m) < \log m - \log \mu(m) \frac{\log \dfrac{1}{1-\rho}}{\log 2} + \left(\frac{\log (l_0 + 2)}{\log 2} + 2\right) \log \frac{1}{1-\rho} + \log 2.$$

That is to say,

$$\frac{\log \dfrac{2}{1-\rho}}{\log 2} \log \mu(m) < \log m + A,$$

where A is a constant depending only on l_0 and ρ, i.e., only on l_0 and θ. We thus have the result stated, namely $\mu(m) < Cm^{\beta}$, where $\beta = \log 2/\log [2/(1-\rho)]$ and C depends on l_0 and θ only.

LEMMA VI. *Let $a = l\theta/(1-\theta)$ $(l > 0)$. Then the number of terms of the sequence lying in the closed interval (a, l) is less than $C(l-a)^{\beta}$, where β and C have the same meaning as in the preceding lemma.*

For let u be the first term of the sequence greater than or equal to a, and v the last term of the sequence less than or equal to l. The number of terms in (a, l) is $\mu(v - u)$. By Lemma V

$$\mu(v - u) < C(v - u)^{\beta} \leq C(l - a)^{\beta},$$

since $u \geq \theta v/(1-\theta)$.

We now prove Theorem III. We consider the interval $(0, m)$, and denote by $\nu(m)$ the number of terms of a sequence satisfying the hypothesis of Theorem III and lying in that interval.

Let I_s be the interval $(m\theta^s(1-\theta)^{-s}, m\theta^{s-1}(1-\theta)^{1-s})$, $s = 1, 2, \cdots$. The sum of the lengths of these intervals is

$$m \frac{1-2\theta}{1-\theta}\left[1 + \frac{\theta}{1-\theta} + \left(\frac{\theta}{1-\theta}\right)^2 + \cdots\right] = m.$$

On the other hand, by Lemma VI,

$$\nu(m) \leq Cm^\beta \frac{(1-2\theta)^\beta}{(1-\theta)^\beta}\left[1 + \left(\frac{\theta}{1-\theta}\right)^\beta + \left(\frac{\theta}{1-\theta}\right)^{2\beta} + \cdots\right]$$

$$= Cm^\beta \frac{(1-2\theta)^\beta}{(1-\theta)^\beta - \theta^\beta} = C_1 m^\beta,$$

C_1 depending only on l_0 and θ. Hence the density of the sequence is at most of order $m^{-\alpha}$, where (ρ being replaced by $\theta/(1-2\theta)$)

$$\alpha = 1 - \beta = 1 - \log 2/\log \frac{2}{1-\rho} = \log\left(\frac{1-2\theta}{1-3\theta}\right) \Big/ \log 2\left(\frac{1-2\theta}{1-3\theta}\right).$$

6. **Proof of Theorem** IV. We consider a sequence U of the type

$$u(n+1) = ng_0 + \left[\frac{n}{2}\right]g_1 + \cdots + \left[\frac{n}{2^p}\right]g_p + \cdots \qquad (g_0 \geq 1,\ g_k \geq 0).$$

By Lemma II this sequence has the complete gap property with respect to θx $(0 < \theta < \frac{1}{3})$ if for $m \geq m_0$ the condition

$$\gamma_m \geq \theta[u(2^m) + 2\gamma_m]$$

is satisfied. This inequality can be written:

$$(g_0 + g_1 + \cdots + g_{m-1})$$
$$\geq \theta[(2^m g_0 + 2^{m-1}g_1 + \cdots + 2g_{m-1}) + (g_0 + \cdots + g_{m-1})].$$

That is to say

$$g_0 + \cdots + g_{m-1} \geq \frac{\theta}{1-\theta} 2^m\left[g_0 + \frac{g_1}{2} + \cdots + \frac{g_{m-1}}{2^{m-1}}\right]. \tag{6.1}$$

We suppose that a is a positive number greater than 2, and take

$$g_0 = 1, \qquad g_k = [a^k] \qquad (k = 1, 2, \cdots).$$

Then

$$g_0 + \cdots + g_{m-1} \geq 1 + \sum_1^{m-1}(a^k - 1) = \frac{a^m - a}{a-1} - (m-2)$$

$$> \frac{a^m}{a-1} - m, \tag{6.2}$$

$$g_0 + \frac{g_1}{2} + \cdots + \frac{g_{m-1}}{2^{m-1}} \leq 1 + \frac{a}{2} + \cdots + \left(\frac{a}{2}\right)^{m-1} = \frac{\left(\frac{a}{2}\right)^m - 1}{\frac{a}{2} - 1},$$

so that

$$\text{(6.3)}\qquad \frac{\theta}{1-\theta}2^m\left[g_0+\cdots+\frac{g_{m-1}}{2^{m-1}}\right]\le\frac{\theta}{1-\theta}\frac{a^m}{\frac{a}{2}-1}-\frac{\theta}{1-\theta}\frac{2^m}{\frac{a}{2}-1}.$$

If we take

$$\frac{1}{a-1}=\frac{\theta}{1-\theta}\frac{1}{\frac{a}{2}-1},$$

that is to say

$$a=\frac{2(1-2\theta)}{1-3\theta}>2 \qquad (0<\theta<\tfrac{1}{3}),$$

we have from (6.2) and (6.3) that (6.1) is satisfied for m greater than some m_0.

Finally

$$u(2^n)<2^n\left[g_0+\frac{g_1}{2}+\cdots+\frac{g_{n-1}}{2^{n-1}}\right]<\frac{a^n}{\frac{a}{2}-1}.$$

Hence if

$$\frac{a^n}{\frac{a}{2}-1}\le m<\frac{a^{n+1}}{\frac{a}{2}-1},$$

we have $\nu(m)\ge 2^n$, where $\nu(m)$ is the number of terms of the sequence not exceeding m. But

$$n+1>\frac{\log m+\log\left(\frac{a}{2}-1\right)}{\log a},$$

and so

$$\frac{\log\nu(m)}{\log 2}+1>\frac{\log m}{\log a}+\frac{\log\left(\frac{a}{2}-1\right)}{\log a}.$$

That is to say

$$\nu(m)\ge Cm^{\log 2/\log a}=Cm^{1-\alpha},$$

where

$$\alpha=1-\frac{\log 2}{\log a}=\log\left(\frac{1-2\theta}{1-3\theta}\right)\Big/\log 2\left(\frac{1-2\theta}{1-3\theta}\right)$$

and C depends only on θ.

As application of Theorems III and IV we notice that any sequence having the complete gap property with respect to the function $\frac{1}{4}x$ has a density of order $m^{-\frac{1}{2}}$ at most, and this order can actually be attained.

7. **Proof of Theorem** V. We begin with two lemmas, in which we treat the cases $\theta > \frac{1}{3}$ and $\theta = \frac{1}{3}$ respectively.

LEMMA VII. *Let u and v $(u < v)$ be any two points of a sequence having the complete gap property with respect to θx $(\frac{1}{3} < \theta < \frac{1}{2})$, and suppose that $v - u \geq l_0$. Then*

$$v \geq \frac{1 - \theta}{1 - 2\theta} u.$$

For suppose that $v < (1 - \theta)u/(1 - 2\theta)$. Then

$$v - u < \left(\frac{1 - \theta}{1 - 2\theta} - 1\right)u = \frac{\theta}{1 - 2\theta} u,$$

that is

$$\left(\frac{1 - 2\theta}{\theta}\right)(v - u) < u.$$

Since $\theta > \frac{1}{3}$ we can find a number θ', $\frac{1}{3} < \theta' < \theta$, such that

$$\left(\frac{1 - 2\theta'}{\theta'}\right)(v - u) < u.$$

θ' being fixed in this way, we consider the interval whose end points are $u - (1 - 2\theta')(v - u)/\theta' > 0$ and $v + (v - u)$. This interval has a length greater than l_0, and the largest gap contained in it cannot exceed $v - u$, since $\theta' > \frac{1}{3}$ and so $(1 - 2\theta')/\theta' < 1$. Hence

$$v - u \geq \theta\left[2(v - u) + \frac{1 - 2\theta'}{\theta'}(v - u)\right],$$

that is,

$$1 \geq \theta\left[2 + \frac{1 - 2\theta'}{\theta'}\right] = \frac{\theta}{\theta'},$$

a result which contradicts $\theta' < \theta$ and so proves the lemma.

The lemma shows that the sequence $u_1 = 0, u_2, \cdots, u_n, \cdots$ consists of terms $U_1, U_2, \cdots, U_q, \cdots$ $(U_1 = u_2)$ satisfying

$$U_q \geq \frac{1 - \theta}{1 - 2\theta} U_{q-1}$$

together with terms which are all included in the open intervals $U_q < x < U_q + l_0$.

LEMMA VIII. *Let u, v, w $(u < v < w)$ be any three points of a sequence having the complete gap property with respect to $\frac{1}{3}x$, and suppose that $w - u \geq l_0$. Then $w \geq 2u$.*

For suppose that $w < 2u$. Then $w - u < u$, and so a fortiori $w - u - 1 < u$. We consider the interval whose end points are $u - (w - u - 1) > 0$ and $w + (w - u - 1)$. This interval has a length greater than l_0, and the largest gap contained in it cannot exceed $w - u - 1$ since the interval (u, w) contains at least the point v. Hence

$$w - u - 1 \geq \tfrac{1}{3}[(w - u) + 2(w - u - 1)] = w - u - \tfrac{2}{3},$$

a relation which is impossible.

The lemma shows that in the case $\theta = \frac{1}{3}$ the sequence $u_1 = 0, u_2, \cdots, u_n, \cdots$ consists of terms $U_1, U_2, \cdots, U_q, \cdots$ $(U_1 = u_2)$ satisfying

$$U_q \geq 2U_{q-1} = \frac{1 - \theta}{1 - 2\theta} U_{q-1},$$

together with (1) terms which, as in the preceding case, are included in the open intervals $U_q < x < U_q + l_0$; (2) one additional term which may be elsewhere in the interval (U_q, U_{q+1}).

We now prove Theorem V. Collecting the results of our two last lemmas we see that if $u_1 = 0, u_2, \cdots, u_n, \cdots$ is a sequence having the complete gap property with respect to θx $(\frac{1}{3} \leq \theta < \frac{1}{2})$ and if $U_q \leq m < U_{q+1}$, then the number $\nu(m)$ of terms of the sequence not exceeding m is at most $1 + 2q + ql_0$. But

$$U_q \geq \left(\frac{1 - \theta}{1 - 2\theta}\right)^{q-1},$$

that is to say,

$$q - 1 \leq \frac{\log U_q}{\log\left(\frac{1 - \theta}{1 - 2\theta}\right)} \leq \frac{\log m}{\log\left(\frac{1 - \theta}{1 - 2\theta}\right)},$$

and so

$$\nu(m) \leq (l_0 + 2)\frac{\log m}{\log\left(\frac{1 - \theta}{1 - 2\theta}\right)} + (l_0 + 3),$$

$$\frac{\nu(m)}{m} \leq \frac{A}{\log\left(\frac{1 - \theta}{1 - 2\theta}\right)} \frac{\log m}{m} + \frac{B}{m},$$

where the numbers A, B depend only on l_0 and not on θ. These values for A, B are not the best possible.

8. **Proof of Theorem** VI. We prove first two lemmas.

LEMMA IX. *Let $u_1, u_2, \cdots, u_n, \cdots$ be an increasing convex sequence of integers; that is, the gaps $u_{n+1} - u_n = \varphi(n)$ are non-decreasing. Suppose that $u_n \leq l < u_{n+1}$ and write $l = u_n + x$ $(0 \leq x < \varphi(n))$. Then in every interval $(a, a + l)$ $(a \geq 0)$ there is a gap of length not less than*

$$G(l) = \min\{\max(x, \varphi(n-1)), \tfrac{1}{2}(u_n + x)\}.$$

(1) If a, $a + l$ lie in the same gap of the sequence, there is nothing to prove since $l = u_n + x > G(l)$.

(2) If a, $a + l$ are in two consecutive gaps of the sequence, there is in $(a, a + l)$ a gap at least equal to $\frac{1}{2}l = \frac{1}{2}(u_n + x) \geq G(l)$.

(3) Suppose that $(a, a + l)$ contains an interval (u_s, u_{s+1}). If $a \leq u_{n-1}$, $(a, a + l)$ necessarily contains a gap of length not less than $\varphi(n - 1)$ and a gap not less than $\min(a + l - u_n, \varphi(n)) \geq \min(l - u_n, \varphi(n)) = \min(x, \varphi(n)) = x$ and therefore a gap of length not less than $\max(x, \varphi(n - 1)) \geq G(l)$. If $a > u_{n-1}$, then $s \geq n$ and so $(a, a + l)$ contains a gap of length not less than $\varphi(n) \geq \max(x, \varphi(n - 1)) \geq G(l)$.

Consideration of the cases (2) and (3) (with $a = 0$) shows that the value of $G(l)$ cannot be improved.

LEMMA X. *Suppose that $\omega(x)$ is a non-negative, differentiable, non-decreasing function for $x \geq 0$ and that $\omega(x) \leq \frac{1}{2}x$, $\omega'(x) \leq 1$ for $x \geq x_0$. Then a necessary and sufficient condition that the convex sequence defined in the preceding lemma should have the complete gap property with respect to $\omega(x)$ is that for $n \geq n_0$*

$$\varphi(n-1) \geq \omega(u_n + \varphi(n-1)). \tag{8.1}$$

The reason for the restriction $\omega(x) \leq \frac{1}{2}x$ will appear below in the proof of Theorem VII.

The condition is necessary by Lemma IX with $x = \varphi(n - 1)$. To prove that it is sufficient we have to show that $\frac{1}{2}(u_n + x) \geq \omega(u_n + x)$, a relation which results from our hypothesis; and also that

$$x \geq \omega(u_n + x) \tag{8.2}$$

when $x > \varphi(n - 1)$. But (8.2) is an immediate consequence of (8.1) and the hypothesis that $\omega'(y) \leq 1$ for y large.

We now prove Theorem VI. We suppose that $\frac{1}{3} \le \theta < \frac{1}{2}$, and take the sequence defined by

$$u_n = \left[\left(\frac{1-\theta}{1-2\theta}\right)^n\right] - n, \quad n \ge 2; \quad u_1 = 0,$$

where $[z]$ denotes the integral part of z. The sequence is convex since $(1-\theta)/(1-2\theta) \ge 2$ and so

$$\begin{aligned} u_{n+1} - 2u_n + u_{n-1} &\ge \left(\frac{1-\theta}{1-2\theta}\right)^{n+1} + \left(\frac{1-\theta}{1-2\theta}\right)^{n-1} - 2 - 2\left(\frac{1-\theta}{1-2\theta}\right)^n \\ &= \left(\frac{1-\theta}{1-2\theta}\right)^{n-1}\left[\frac{1-\theta}{1-2\theta} - 1\right]^2 - 2 \ge 0, \end{aligned}$$

for $n \ge 3$. Similarly $u_3 - 2u_2 \ge 0$.

Moreover, the sequence has the complete gap property with respect to θx. In fact, by Lemma X, it is sufficient to prove that

$$u_n - u_{n-1} \ge \theta(2u_n - u_{n-1}),$$

that is

$$u_n(1-2\theta) \ge u_{n-1}(1-\theta).$$

But we have for $n > 1$

$$u_n > \left(\frac{1-\theta}{1-2\theta}\right)^n - 1 - n, \qquad u_{n-1} \le \left(\frac{1-\theta}{1-2\theta}\right)^{n-1} - (n-1),$$

and so it is sufficient to prove that

$$\left(\frac{1-\theta}{1-2\theta}\right)^n - 1 - n \ge \left(\frac{1-\theta}{1-2\theta}\right)^n - (n-1)\left(\frac{1-\theta}{1-2\theta}\right),$$

or

$$(n-1)\left(\frac{1-\theta}{1-2\theta}\right) \ge n + 1,$$

a relation which is true for $n \ge 3$ since $(1-\theta)/(1-2\theta) \ge 2$.

It remains to find the order of the density of the sequence. Suppose that $u_n \le m < u_{n+1}$. Then $\nu(m) = n$, and since

$$u_{n+1} \le \left(\frac{1-\theta}{1-2\theta}\right)^{n+1} - (n+1),$$

we have

$$\log u_{n+1} \leq (n + 1) \log \left(\frac{1 - \theta}{1 - 2\theta}\right).$$

Hence

$$\nu(m) \geq \frac{\log m}{\log \left(\dfrac{1 - \theta}{1 - 2\theta}\right)} - 1,$$

$$\frac{\nu(m)}{m} \geq \frac{1}{\log \left(\dfrac{1 - \theta}{1 - 2\theta}\right)} \frac{\log m}{m} - \frac{1}{m}.$$

This completes the proof of Theorem VI.

9. **Proof of Theorem** VII. This theorem is trivial and we give it only for completeness.

Let u, v $(u \neq 0)$, $(u < v)$ be two points of the sequence, and consider the interval $(u - 1, 2v - u)$. The maximum gap in this interval is $v - u$, whereas the length of the interval is $2(v - u) + 1$. Thus if $\omega(x) \geq \frac{1}{2}x$, we must have $2(v - u) + 1 < l_0$, that is to say, the sequence is finite.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY AND STANFORD UNIVERSITY.

AFFINE GEOMETRY OF VECTOR SPACES OVER RINGS

By C. J. Everett

1. **Introduction.** The intuitive ideas involved in the "geometry" of the points with integer coördinates (lattice points) in the affine real plane and the subsets (lines) of such points which are in alignment are here abstracted to their essential content, in the spirit and with the methods of E. Artin's investigation on affine geometry [1]. The unusual character of such a "plane" geometry arises from the fact that parallel lines cannot be defined simply as lines which do not intersect. It is curious that parallelism may be introduced as an abstract equivalence relation with certain natural properties.

In §2, a geometry of points and lines is defined in a 2-space over a general type of ring and is shown to have certain geometric properties I-VII. These properties are then adopted as axioms for an abstract geometry of undefined points and lines in §3 and are proved by introduction of coördinates to completely characterize the 2-space geometry.

2. **Affine geometry in a 2-space over a ring.** Let $\mathfrak{R} = (0, 1, \alpha, \beta, \cdots)$ be a ring with unit ($1\alpha = \alpha = \alpha 1$) and no zero divisors ($\alpha\beta = 0$ implies $\alpha = 0$ or $\beta = 0$). If $\alpha = \beta\gamma$ we say γ divides α. There will be no occasion to consider left divisors. Suppose further that every two elements α, β, not both zero, have a g.c.d., i.e., there is a δ which divides α, β, and which is divisible by every δ' dividing α, β.

A number of properties of such a ring, embodied in the following lemmas, will be needed.

Lemma 1. *If $\delta_2\delta_1 = 1$, then $\delta_1\delta_2 = 1$ and δ_1, δ_2 are units.*

For $\delta_2(\delta_1\delta_2 - 1) = (\delta_2\delta_1)\delta_2 - \delta_2 = \delta_2 - \delta_2 = 0$. Since $\delta_2 \neq 0$, $\delta_1\delta_2 = 1$.

Lemma 2. *If δ is a g.c.d. of α, β, then $\delta\gamma$ is a g.c.d. of $\alpha\gamma$, $\beta\gamma$.*

Let δ' be a g.c.d. of $\alpha\gamma$, $\beta\gamma$, and write $\alpha = \alpha_1\delta$, $\beta = \beta_1\delta$, $\alpha\gamma = \alpha_2\delta'$, $\beta\gamma = \beta_2\delta'$. Since $\alpha\gamma = \alpha_1\delta\gamma$, $\beta = \beta_1\delta\gamma$, $\delta\gamma$ divides δ' and $\delta' = \delta_1\delta\gamma$. Thus $\alpha\gamma = \alpha_2\delta_1\delta\gamma$, $\beta\gamma = \beta_2\delta_1\delta\gamma$, and therefore $\delta_1\delta$ divides α, β. Hence $\delta = \delta_2\delta_1\delta$, $\delta_2\delta_1 = 1$, and $\delta' = \omega\delta\gamma$, where $\omega = \delta_1$ is a unit. Every common divisor of $\alpha\gamma$, $\beta\gamma$ divides δ' and hence $\delta\gamma$.

Lemma 3. *Relations $\alpha_1\delta = \alpha_2\delta'$, $\beta_1\delta = \beta_2\delta'$, where none of the pairs (α_1, β_1), (α_2, β_2), (δ, δ') is the zero pair $(0, 0)$, and where the former two pairs are relatively prime (g.c.d. $= 1$), imply $\alpha_2 = \alpha_1\omega$, $\beta_2 = \beta_1\omega$, where ω is a unit of $\mathfrak{R}$.*

Received May 20, 1942.

Since δ' divides $\alpha_1\delta$, $\beta_1\delta$, δ' divides δ (Lemma 2). Similarly δ divides δ'. Hence $\delta = \omega_1\delta'$, $\delta' = \omega_2\delta$. If (say) $\delta \neq 0$, $\delta = \omega_1\omega_2\delta$, $\delta(1 - \omega_1\omega_2) = 0$ and ω_1 is a unit. Hence $\alpha_2\delta' = \alpha_1\delta = \alpha_1\omega_1\delta'$, $\alpha_2 = \alpha_1\omega_1$. Similarly $\beta_2 = \beta_1\omega_1$.

For such a ring $\mathfrak{R}$ consider the vector space $V(\mathfrak{R})$ of all elements (α, β), α, $\beta \in \mathfrak{R}$. By $AG(V(\mathfrak{R}))$ is meant the geometry consisting of the elements as points, and of the sets of points (called lines) of type $(\alpha + \mu\kappa, \beta + \nu\kappa)$, where κ is arbitrary in $\mathfrak{R}$, α, β fixed, and μ, ν not both zero, fixed, and relatively prime. The lines $(\alpha + \mu\kappa, \beta + \nu\kappa)$ and $(\alpha' + \mu'\kappa, \beta' + \nu'\kappa)$ are identical (consist of the same points) if and only if, for some $\kappa \in \mathfrak{R}$, and some unit ω, $\alpha' = \alpha + \mu\kappa$, $\beta' = \beta + \nu\kappa$ and $\mu' = \mu\omega$, $\nu' = \nu\omega$. This is an immediate consequence of Lemma 3. Hence, if (α', β') belongs to the line $(\alpha + \mu\kappa, \beta + \nu\kappa)$, the latter may be written $(\alpha' + \mu\kappa, \beta' + \nu\kappa)$.

It is natural to say a point is on a line in case it is in the point set of the line. Now define the relation $(\alpha_1 + \mu_1\kappa, \beta_1 + \nu_1\kappa) \parallel (\alpha_2 + \mu_2\kappa, \beta_2 + \nu_2\kappa)$ on lines (parallelism) to mean there is a unit $\omega \in \mathfrak{R}$ such that $\mu_2 = \mu_1\omega$, $\nu_2 = \nu_1\omega$. (I) This is determinative, reflexive, transitive.

It is elementary to check that (II) every line $(\alpha + \kappa, \beta)$ cuts every line $(\alpha', \beta' + \kappa)$; (III) (α, β) and $(\alpha + \mu, \beta + \nu)$ are distinct points on line $(\alpha + \mu\kappa, \beta + \nu\kappa)$; (IV) every two distinct points (α_1, β_1), (α_2, β_2) are on one, and only one, (Lemma 3) line: $(\alpha_1 + \mu\kappa, \beta_1 + \nu\kappa)$, where $\alpha_2 - \alpha_1 = \mu\delta$, $\beta_2 - \beta_1 = \nu\delta$, $\delta =$ g.c.d. $\alpha_2 - \alpha_1$, $\beta_2 - \beta_1$; (V) for every point (α_1, β_1) and line $(\alpha + \mu\kappa, \beta + \nu\kappa)$ there is exactly one line $(\alpha_1 + \mu\kappa, \beta_1 + \nu\kappa)$ on the point and parallel to the line.

A *translation* in this geometry is defined as a point-to-point mapping $\sigma(P) = P'$ which is either the identity, or which (a) is one-one, (b) has no fixed point, (c) preserves parallel lines: $P \cup Q \parallel \sigma(P) \cup \sigma(Q)$, all P, $Q (P \cup Q$, the unique line on P, $Q)$, (d) has parallel "traces": $P \cup \sigma(P) \parallel Q \cup \sigma(Q)$.

Obviously $\sigma(\alpha, \beta) = (\alpha + \mu, \beta + \nu)$, μ, ν fixed, is a translation. Conversely, every translation σ is of this type. For, let $\sigma(0, 0) = (\mu, \nu) \neq (0, 0)$, $\mu = \mu_1\delta$, $\nu = \nu_1\delta$, $\delta =$ g.c.d. μ, ν, and let (α, β) be any point not on the line $(\mu_1\kappa, \nu_1\kappa)$, with $\alpha = \alpha_1\delta_1$, $\beta = \beta_1\delta_1$, $\delta_1 =$ g.c.d. α, β. If $\sigma(\alpha, \beta) = (\alpha', \beta')$, then by (c), $\alpha' - \mu = \alpha_1\delta_2$, $\beta' - \nu = \beta_1\delta_2$, and by (d), $\alpha' - \alpha = \mu_1\delta_3$, $\beta' - \beta = \nu_1\delta_3$. From the latter one has $\alpha_1(\delta_1 - \delta_2) = \mu_1(\delta - \delta_3)$, $\beta_1(\delta_1 - \delta_2) = \nu_1(\delta - \delta_3)$. By Lemma 3, $\alpha' = \alpha + \mu$, $\beta' = \beta + \nu$. A similar argument establishes the result for (α, β) on line $(\mu_1\kappa, \nu_1\kappa)$ using instead of $(0, 0)$ a point not on $(\mu_1\kappa, \nu_1\kappa)$.

Thus (VI) for every two points P, Q there is a translation taking P into Q.

Since the correspondence $\tau \to [\alpha, \beta]$ is one-one from the set T of all translations to a complete 2-space of pairs $[\alpha, \beta]$ over $\mathfrak{R}$, where $\tau(0, 0) = (\alpha, \beta)$, and $\tau\tau' \to [\alpha + \alpha', \beta + \beta']$ when $\tau'(0, 0) = (\alpha', \beta')$, it follows that T is isomorphic to the 2-space over $\mathfrak{R}$.

Finally, we shall establish a fundamental property (VII) of the endomorphisms of T by means of the following lemma, which seems to have intrinsic interest.

Two vectors $[\alpha_1, \beta_1]$, $[\alpha_2, \beta_2]$ over $\mathfrak{R}$ are said to have the same *slope* in case $\alpha_1 = \alpha_1'\delta_1$, $\beta_1 = \beta_1'\delta_1$, $\alpha_2 = \alpha_2'\delta_2$, $\beta_2 = \beta_2'\delta_2$, $\delta_i =$ g.c.d. α_i, β_i imply $\alpha_2' = \alpha_1'\omega$, $\beta_2' = \beta_1'\omega$ for some unit $\omega \in \mathfrak{R}$. An endomorphism of T is any mapping θ of T into T such that $([\alpha, \beta] + [\gamma, \delta])\theta = [\alpha, \beta]\theta + [\gamma, \delta]\theta$.

LEMMA 4. *An endomorphism of T such that $[\alpha, \beta]\theta = [\alpha', \beta']$ has the same slope as $[\alpha, \beta]$ whenever $[\alpha', \beta'] \neq [0, 0]$ is defined by $[\alpha, \beta]\theta = [\alpha\gamma, \beta\gamma]$ for some fixed $\gamma \in \mathfrak{R}$, and conversely.*

Let $[1, 0]\theta = [\gamma_1, 0]$, $[0, 1]\theta = [0, \gamma_2]$. Then $[1, 1]\theta = [\gamma, \gamma] = [1, 0]\theta + [0, 1]\theta = [\gamma_1, \gamma_2]$, whence $\gamma_1 = \gamma_2 = \gamma$. Since $[\alpha, 1]\theta = [\alpha, 0]\theta + [0, 1]\theta = [\alpha', 0] + [0, \gamma] = [\alpha', \gamma]$ and $[\alpha, 1]\theta = [\alpha\mu, \mu]$, then $\mu = \gamma$ and $\alpha' = \alpha\gamma$, $[\alpha, 0]\theta = [\alpha\gamma, 0]$. Finally $[\alpha, \beta]\theta = [\alpha, 0]\theta + [0, \beta]\theta = [\alpha\gamma, \beta\gamma]$.

For every set of parallel lines having the form $(\alpha + \mu\kappa, \beta + \nu\kappa)$ every translation $\tau \neq 0$ with these lines as traces must have the form $[\mu\gamma, \nu\gamma]$, since $\tau(0, 0) = (\mu\gamma, \nu\gamma)$ for some $\gamma \in \mathfrak{R}$. Thus (VII) for every class of parallel lines, there is a translation $[\mu, \nu]$ such that every translation with this class of lines as traces is of the form $[\mu, \nu]\theta$ for some slope preserving endomorphism θ of T.

3. **Axioms for a lattice point geometry.**

In this section is developed a set of purely geometric axioms completely characterizing $AG(V(\mathfrak{R}))$, following Artin [1]. Consider two classes of objects: points $P, Q, R, \cdots$, and lines $l, m, n, \cdots$, with two undefined relations: $P \in l$ and $l \parallel m$ (P on l, l parallel to m), subject to the axioms:

AXIOM I. *Parallelism is an equivalence relation throwing all lines into disjoint classes $\mathfrak{L}$: $l \parallel l$; $l \parallel m$ implies $m \parallel l$; $l \parallel m$ and $m \parallel n$ implies $l \parallel n$.*

AXIOM II. *There is a parallel class $\mathfrak{L}_1$ and a line $l_2 \in' \mathfrak{L}_1$ such that for every line $l_1 \in \mathfrak{L}_1$ there is a point P on l_1 and l_2. There is a line cutting every line of the class $\mathfrak{L}_1$.*

AXIOM III. *There are at least two points on each line.*

AXIOM IV. *For every two distinct points P, Q there is a unique line $P \cup Q$ with P, Q thereon.*

AXIOM V. *For every point P and line l there is exactly one line $m \parallel l$ with P on m.*

DEFINITION. A mapping σ of all the points P into a subset: $\sigma(P) = P'$ is called a *dilatation* in case for every two distinct points $P \neq Q$ there is a line $l \parallel P \cup Q$ with $\sigma(P)$ and $\sigma(Q)$ on l.

THEOREM 1. *A dilatation σ is completely determined by images of any pair of distinct points. Either $\sigma(P)$ is constant: $\sigma(P) = Q$ for all P, or σ is one-one: $P \neq Q$ implies $\sigma(P) \neq \sigma(Q)$.*

For the proof, see [1; Theorem 1].

Remark. One-one dilatations need not have inverses; e.g., $\sigma(\alpha, \beta) = (2\alpha, 2\beta)$ in $AG(V(C))$, where C is the ring of rational integers.

COROLLARY 1. *A dilatation has either all points fixed (identity $1(P) = P$, all P), or one point fixed, or no point fixed.*

From this point, all dilatations are supposed one-one.

Definition. If σ is a dilatation, any line l with P and $\sigma(P)$ on l is called a *σ-trace* of P. If P is not a fixed point, it has a unique trace.

Theorem 2. *If Q is on l, a σ-trace of P, then $\sigma(Q)$ is on l.*

Corollary 2. *If $\sigma(P) = P$, $\sigma \neq 1$, every σ-trace contains P.*

Corollary 3. *A point Q on two distinct σ-traces is fixed.*

Definition. If $\sigma = 1$, or if σ has all traces parallel (and hence no fixed point), σ is called a *translation*.

Remark. A dilatation $\sigma \neq 1$ with no fixed point need not have parallel traces; e.g., $\sigma(\alpha, \beta) = (1 - \alpha, 1 - \beta)$ in $AG(V(C))$, where the C are rational integers.

Axiom VI. *For every pair of points $P \neq Q$, there is a translation τ_{PQ} such that $\tau_{PQ}(P) = (Q)$.*

Corollary 4. *A translation $\neq 1$ is completely determined by the image of one point. Hence the translation of the above axiom is unique.*

If $\tau(A) = B$, $C \varepsilon' A \cup B$, then $\tau(C)$ is the unique point common to l on $B \parallel A \cup C$ and m on $C \parallel A \cup B$. But then $\tau(D)$ for $D \varepsilon A \cup B$ is uniquely determined by the same argument since $D \varepsilon' C \cup \tau(C)$.

Corollary 5. *If $C \varepsilon l \parallel A \cup B$ and $B \varepsilon m \parallel A \cup C$, then l and m have a common point. A parallelogram with three vertices has a fourth.*

Since the translation τ_{AC} takes A into C, $\tau_{AC}(B)$ is on l and m.

Theorem 3. *The set T of all translations is an Abelian group.*

If τ_1 and τ_2 are translations, $\tau_1\tau_2$ is also. We may suppose that neither τ_1 nor $\tau_2 = 1$. If $\tau_1\tau_2(P) = P$, then τ_1 and τ_2 have the same class of traces. If $Q \varepsilon' P \cup \tau_2(P)$ and $\tau_1\tau_2(Q) \neq Q$, then $Q \cup \tau_1\tau_2(Q) \parallel P \cup \tau_2(P)$. But $P \varepsilon \tau_1\tau_2(Q) \cup Q$, whence $Q \cup \tau_1\tau_2(Q) = P \cup \tau_2(P)$. The contradiction implies $\tau_1\tau_2(Q) = Q$ and $\tau_1\tau_2 = 1$. Now if τ_1, τ_2 have the same traces, $\tau_1\tau_2$ must have parallel traces. If τ_1, τ_2 have different traces, then $\tau_1\tau_2 = \tau_2\tau_1$ [1; Theorem 5]. Suppose τ_1, τ_2 have different traces. We prove that, for $P \neq Q$, $P \cup \tau_1\tau_2(P) \parallel Q \cup \tau_1\tau_2(Q)$.

Case 1. $Q \varepsilon' P \cup \tau_1(P)$, $Q \varepsilon' P \cup \tau_2(P)$. Then $P \cup Q \varepsilon'$ trace class of τ_1, $P \cup Q \varepsilon'$ trace class of τ_2. Hence if τ_{PQ} is the $P \to Q$ translation, $\tau_{PQ}\tau_1\tau_2(P) = \tau_1\tau_2(Q)$, and $P \cup \tau_1\tau_2(P) \parallel Q \cup \tau_1\tau_2(Q)$.

Case 2. $Q = \tau_1(P)$ or $Q = \tau_2(P)$. Suppose the former. Then $\tau_2(Q) = \tau_2\tau_1(P) = \tau_1\tau_2(P)$, and $\tau_1(\tau_2 Q) \cup Q \parallel P \cup \tau_1\tau_2(P)$. Also $\tau_1(\tau_2 Q) \cup Q = \tau_1(\tau_1\tau_2 P) \cup \tau_1(P) \parallel \tau_1\tau_2(P) \cup P$.

Case 3. $Q \varepsilon P \cup \tau_1(P)$ or $Q \varepsilon P \cup \tau_2(P)$, but $Q \neq \tau_1(P)$ and $Q \neq \tau_2(P)$. Let $P' = \tau_1\tau_2(P)$. Then $Q \varepsilon' P' \cup \tau_1 P'$ and $Q \varepsilon' P' \cup \tau_2 P'$. By case 1, $P \cup \tau_1\tau_2(P) \parallel P' \cup \tau_1\tau_2 P' \parallel Q \cup \tau_1\tau_2 Q$. This completes the proof that T is closed under multiplication; hence we have the "little" Desargues theorem.

If $\tau \varepsilon T$, then τ^{-1} exists and is a translation. For, let $Q' \varepsilon' P \cup \tau(P)$. By Corollary 5, lines l on $P \parallel \tau(P) \cup Q'$ and m on $Q' \parallel P \cup \tau(P)$ intersect in a unique point Q, and clearly $\tau(Q) = Q'$. Obviously τ^{-1} is a translation. The Abelian property now follows as in [1; Theorems 5, 6], since our Axioms II, III provide three non-collinear points.

Definition. Let $\mathfrak{T} = \{0, 1, \alpha, \beta, \cdots\}$ be the ring of all endomorphisms of T, subject to the definitions: $(\tau_1\tau_2)^\alpha = \tau_1^\alpha\tau_2^\alpha$; $\alpha = \beta$ means $\tau^\alpha = \tau^\beta$ for all $\tau \varepsilon T$, $\tau^{\alpha+\beta} = \tau^\alpha\tau^\beta$, $\tau^{\alpha\beta} = (\tau^\alpha)^\beta$, $\tau^0 = 1$ for all $\tau \varepsilon T$, $\tau^1 = \tau$ for all $\tau \varepsilon T$. Let $\mathfrak{R}$ be the subset of all elements α such that the traces of τ are among the traces of τ^α for each $\tau \varepsilon T$.

Lemma 5. *$\mathfrak{R}$ is a ring with unit* 1.

The proof is immediate from the definitions.

Theorem 4. *Let P be any particular point, and let $\alpha \varepsilon \mathfrak{R}$. Then $\sigma(Q) = \tau_{PQ}^\alpha(P)$ is a dilatation, and $\tau^\alpha\sigma = \sigma\tau$ (all $\tau \varepsilon T$). If σ is constant, $\tau^\alpha = 1$ ($\tau \varepsilon T$) and $\alpha = 0$.*

The proof is essentially due to Artin [1; Theorem 8]. One has, for all Q, R, $\tau_{QR}^\alpha\tau_{PQ}^\alpha(P) = \tau_{PR}^\alpha(P)$ and $\tau_{QR}^\alpha\sigma(Q) = \sigma(R)$. Either $\tau_{QR}^\alpha = 1$, $\sigma(Q) = \sigma(R)$, and there is a line on both $\parallel Q \cup R$, or $\tau_{QR}^\alpha \neq 1$, $\sigma(Q) \neq \sigma(R)$ and $\sigma(Q) \cup \sigma(R) \parallel Q \cup R$. Hence σ is a dilatation. Now $\sigma(P) = P$. If $\sigma(Q)$ is constant, $\tau_{QR}^\alpha(P) = P$ (all Q, R), $\tau_{QR}^\alpha = 1$ (all Q, R) and $\tau^\alpha = 1$ ($\tau \varepsilon T$). In general, $\tau_{QR}^\alpha\,\sigma(Q) = \sigma\tau_{QR}(Q)$. Now $\tau_{QR}^\alpha\sigma$ and $\sigma\tau_{QR}$ are dilatations, with the same effect on Q and P. For, $\tau_{QR}^\alpha\sigma(P) = \tau_{QR}^\alpha(P)$. If $S = \tau_{QR}(P)$, then $\sigma\tau_{QR}(P) = \sigma(S) = \tau_{PS}^\alpha(P) = \tau_{QR}^\alpha(P)$, for $\tau_{PS} = \tau_{QR}$, since both have the same effect on P.

Lemma 6. *If for some $\tau_0 \varepsilon T$, $\alpha \varepsilon \mathfrak{R}$, $\tau_0^\alpha = 1$, then $\alpha = 0$ or $\tau_0 = 1$.*

For this $\alpha \varepsilon \mathfrak{R}$, $\tau^\alpha\sigma = \sigma\tau$ (all $\tau \varepsilon T$). Hence $\tau_0^\alpha\sigma = \sigma\tau_0$, and $\sigma(Q) = \sigma\tau_0(Q)$ for all Q. If $\tau_0 \neq 1$, then $Q \neq \tau_0(Q) = Q'$, and $\sigma(Q) = \sigma(Q')$. Hence σ is constant and $\alpha = 0$.

Lemma 7. *In the ring $\mathfrak{R}$, $\alpha\beta = 0$ implies $\alpha = 0$ or $\beta = 0$.*

For, let τ_0 be any translation $\neq 1$, and suppose $\alpha\beta = 0$ with $\alpha \neq 0$, $\beta \neq 0$. Then $\tau_0^\alpha \neq 1$, $(\tau_0^\alpha)^\beta \neq 1$. But $\tau_0^{\alpha\beta} = 1$.

Remark. Every line in $\mathfrak{L}_1$ cuts every line l_2' of the parallel class $\mathfrak{L}_2$ of l_2 of Axiom II. For, l_2 cuts l_1. Let P be any point on l_2'. Draw $l_1' \parallel l_1$ through P. Then l_2 cuts l_1'.

Axiom VII. *For each parallel class $\mathfrak{L}$, there is a unit translation τ_0 with traces in $\mathfrak{L}$, such that for every translation τ with traces in $\mathfrak{L}$, $\tau = \tau_0^\alpha$ for some $\alpha \varepsilon \mathfrak{R}$.*

Lemma 8. *If τ_1, τ_2 are the unit translations of $\mathfrak{L}_1$ and $\mathfrak{L}_2$ of Axiom* II, *then every translation τ is uniquely expressible in the form $\tau = \tau_1^\alpha\tau_2^\beta$.*

Let $P \varepsilon l_2 \varepsilon \mathfrak{L}_2$, $Q = \tau(P) \varepsilon l_1 \varepsilon \mathfrak{L}_1$, and R be the common point of l_1, l_2. Then $\tau_{RQ}\tau_{PR} = \tau_1^\alpha\tau_2^\beta = \tau$. Uniqueness follows from Lemma 6.

We now introduce coördinates with origin O at any fixed point, and τ_1, τ_2 the unit translations of $\mathfrak{P}_1$, $\mathfrak{P}_2$. Let P be any point and write $\tau_{OP} = \tau_1^{\alpha}\tau_2^{\beta}$. Assign to P coördinates (α, β). The correspondence $P \to (\alpha, \beta)$ is obviously one-one from all points onto all pairs over $\mathfrak{R}$.

If l is any line, $P = (\alpha, \beta)$ a fixed point of l, τ_0 the unit translation for the parallel class of l, Q an arbitrary point of l, then $\tau_{OQ} = \tau_{PQ}\tau_{OP} = \tau_0^{\kappa}\tau_{OP} = (\tau_1^{\mu}\tau_2^{\nu})^{\kappa}\tau_1^{\alpha}\tau_2^{\beta}$, so $Q = (\alpha + \mu\kappa, \beta + \nu\kappa)$, with μ, ν not both 0. Clearly every $\kappa \in \mathfrak{R}$ is realized for some $Q \in l$. Let $Q = \tau_0^{\kappa}(P)$. If $\mu = \mu_1\delta$, $\nu = \nu_1\delta$, $\delta \in \mathfrak{R}$, then $\tau_0 = \tau_1^{\mu}\tau_2^{\nu} = (\tau_1^{\mu_1}\tau_2^{\nu_1})^{\delta}$. The trace class of (τ_0) is that of $\tau_1^{\mu_1}\tau_2^{\nu_1}$; hence the latter translation has the form τ_0^{ρ}, $\rho \in \mathfrak{R}$. Thus $\mu_1 = \mu_1\delta\rho$, $\nu_1 = \nu_1\delta\rho$, $\delta\rho = 1$, and μ, ν are "relatively prime".

Finally, let ξ, η be any pair of elements of $\mathfrak{R}$ not both 0, and consider the line $l = (\mu\kappa, \nu\kappa)$ through $O = (0, 0)$ and $P = (\xi, \eta)$. Then $\xi = \mu\kappa$, $\eta = \nu\kappa$, and κ is a common right divisor of ξ, η. Moreover, if $\xi = \xi_1\delta$, $\eta = \eta_1\delta$, then the trace class of $\tau_1^{\xi_1}\tau_2^{\eta_1}$ is that of the unit translation τ_0 for the class of l, and the former translation is τ_0^{ρ}, $\rho \in \mathfrak{R}$. Thus $\xi_1 = \mu\rho$, $\eta_1 = \nu\rho$, $\xi = \xi_1\delta = \mu\rho\delta = \mu\kappa$, $\eta = \eta_1\delta = \nu\rho\delta = \nu\kappa$, δ divides κ, and κ is a g.c.d. of ξ, η.

THEOREM 5. *A geometry satisfying Axioms* I—VII *admits a coördinate system in which $P = (\alpha, \beta)$ with α, β in a ring $\mathfrak{R}$ with unit, without zero-divisors, and with every pair of elements not both zero possessing a g.c.r.d. The points on a line have the form $(\alpha + \mu\kappa, \beta + \nu\kappa)$ where α, β are fixed, where μ, ν are fixed and not both* 0, *and have g.c.d.* 1, *and where κ is arbitrary in $\mathfrak{R}$.*

4. **Completely intersecting parallel classes.** We have seen how, in $AG(V(\mathfrak{R}))$, every line of a class $\mathfrak{P}$ is of the form $(\alpha + \mu\kappa, \beta + \nu\kappa)$ with μ, ν fixed for $\mathfrak{P}$ and g.c.d. $\mu, \nu = 1$. Axioms II, VI provide classes of "completely intersecting" lines: every l_1 of $\mathfrak{P}_1$ cuts every l_2 of $\mathfrak{P}_2$.

THEOREM 6. *The classes of lines $(\alpha_1 + \alpha_{11}\kappa, \beta_1 + \alpha_{21}\kappa)$, $(\alpha_2 + \alpha_{12}\kappa, \beta_2 + \alpha_{22}\kappa)$ are completely intersecting if and only if the matrix $A = (\alpha_{ij})$ has a right inverse: $AB = I$.*

If $AB = I$, the arbitrary pair of lines above intersect, since $B\begin{pmatrix}\alpha_2 - \alpha_1\\ \beta_2 - \beta_1\end{pmatrix}$ provides a solution $\begin{pmatrix}\kappa\\ -\lambda\end{pmatrix}$ for the equations $\alpha_1 + \alpha_{11}\kappa = \alpha_2 + \alpha_{12}\lambda$, $\beta_1 + \alpha_{21}\kappa = \beta_2 + \alpha_{22}\lambda$. Conversely, if the classes are completely intersecting, the inverse exists by intersection of $(\alpha_{11}\kappa, \alpha_{21}\kappa)$ with $(1 + \alpha_{12}\kappa, \alpha_{22}\kappa)$ and $(\alpha_{11}\kappa, -1 + \alpha_{21}\kappa)$ with $(\alpha_{12}\kappa, \alpha_{22}\kappa)$.

BIBLIOGRAPHY

1. E. ARTIN, *Coördinates in Affine Geometry*, Reports of Math. Colloquium, Notre Dame University Press, Ser. 2, Issue 2, 1940, pp. 15–20.

UNIVERSITY OF MICHIGAN.

A THEOREM OF CARTWRIGHT

By R. P. Boas, Jr. and A. C. Schaeffer

An important theorem of Cartwright [1], [2], [6], [7] states:

Let $f(z)$ be an entire function of exponential type, satisfying

$$f(z) = O(e^{k|z|}) \qquad (|z| \to \infty; 0 < k < \pi),$$

$$|f(n)| \le M \qquad (n = 0, \pm 1, \pm 2, \cdots).$$

Then $f(z)$ is bounded on the real axis,

$$(1) \qquad |f(x)| \le A(k)M \qquad (-\infty < x < \infty).$$

The theorem becomes false if the condition $k < \pi$ is omitted, as is shown by the example $f(z) = z \sin \pi z$. It then appears plausible that if k approaches π the constant $A(k)$ in (1) must become infinite. This is indeed the case; the purpose of this note is to show that the order of $A(k)$ is precisely $\log \{1/(\pi - k)\}$. Our result is:

If $A(k)$ is the least possible constant for which Cartwright's theorem is true, then

$$\frac{2}{\pi} \log \left\{ \frac{4}{\pi} \left[\frac{\pi}{\pi - k} \right] \right\} \le A(k) \le 4 + 2e \log \frac{\pi}{\pi - k}.$$

The only previous upper estimates of $A(k)$ seem to be due to Macintyre [6] and Boas [1]. They obtained the inequalities

$$A(k) \le \frac{1}{\pi^2(\pi - k)} \left\{ \tfrac{2}{3}(\pi - k)^2 + 2\pi^2(\pi - k) + \frac{\pi^3}{3} + \sum_1^\infty \frac{2}{n^3} \right\}$$

and

$$A(k) \le 2 + \frac{\pi}{3(\pi - k)}$$

respectively.

Cartwright's theorem may be thought of as a generalization of a theorem of Landau [5], Gronwall [4] and Grandjot [3], which states that a polynomial of degree $r - 1$ $(r \ge 2)$, bounded by M at the r-th roots of unity, is bounded in the unit circle by $B(r)M$, where $B(r)$ depends only on r. In fact, an equivalent statement is that a trigonometric sum

$$T(\theta) = \sum_{j=0}^{r-1} c_j e^{ij\theta},$$

Received May 26, 1942.

where

$$| T(2\pi\nu/r) | \leq M \qquad (\nu = 0, 1, \cdots, r-1),$$

satisfies

$$| T(\theta) | \leq B(r)M \tag{2}$$

for all real θ. If we set

$$f(z) = e^{-i\pi z(r-1)/r} T(2\pi z/r),$$

then $f(z)$ is an entire function of exponential type $\pi(r-1)/r < \pi$, and $| f(\nu) | \leq M$ ($\nu = 0, \pm 1, \pm 2, \cdots$). Gronwall and Grandjot showed by an example that the best possible $B(r)$ in (2) is asymptotic to $(2/\pi) \log r$ as $r \to \infty$. Our example is different from theirs.

1. Proof that

$$A(k) \leq 4 + 2e \log \{\pi/(\pi - k)\}. \tag{3}$$

Given a function $f(z)$ which satisfies the conditions of Cartwright's theorem with $k < \pi$, let $\omega = (\pi - k)/\pi$ and let z_0 be an arbitrary point in the z-plane. The function

$$g(z) = f(z) \frac{\sin \omega(z - z_0)}{\omega(z - z_0)} \tag{4}$$

is an entire function of exponential type $k + \omega < \pi$ and it satisfies

$$\sum_{-\infty}^{\infty}{}' \left| \frac{g(n)}{n} \right| < \infty. \tag{5}$$

An interpolation formula of Pólya and Szegö and Valiron states that, if $g(z)$ is an entire function of exponential type less than π for which (5) is true, then

$$g(z) = \frac{\sin \pi z}{\pi} \sum_{-\infty}^{\infty} \frac{(-1)^n g(n)}{z - n}. \tag{6}$$

(This formula was proved independently and at about the same time by Valiron [11] and Pólya and Szegö [9]. Pólya and Szegö give the formula under somewhat less general conditions than we need here. The formula in full generality is a consequence of the following lemma: *If an entire function $\varphi(z)$ satisfies an inequality of the form $| \varphi(z) | \leq | z |^{a} e^{\pi |z|}$ and vanishes at $z = 0, \pm 1, \pm 2, \cdots$, then $\varphi(z) = P(z) \sin \pi z$, where $P(z)$ is a polynomial.* This lemma was first stated by Pólya [8]. It was rediscovered by Valiron, and forms the basis of his proof of the interpolation formula.)

Thus (4) and (6) give

$$f(z) \frac{\sin \omega(z - z_0)}{\omega(z - z_0)} = \frac{\sin \pi z}{\pi} \sum_{-\infty}^{\infty} \frac{(-1)^n f(n) \sin \omega(z_0 - n)}{\omega(z - n)(z_0 - n)}.$$

Keeping z fixed and letting z_0 approach z, we obtain

$$f(z) = \frac{\sin \pi z}{\pi} \sum_{-\infty}^{\infty} \frac{(-1)^n f(n) \sin \omega(z - n)}{\omega(z - n)^2}. \tag{7}$$

(This is an interpolation formula of Boas [1], here derived in a simpler way.) Since $f(z)$ is bounded by M at the integers, (7) gives the inequality

$$| f(x) | \leq 2M + M\Big(\sum_{-\infty}^{[x]-1} + \sum_{[x]+2}^{\infty}\Big) \left| \frac{\sin \pi(x - n) \sin \omega(x - n)}{\pi\omega(x - n)^2} \right|.$$

Now $| \theta^{-1} \sin \theta | \leq 1$. Hence, if $0 < \epsilon < 1$,

$$\left|\frac{\sin \omega(x - n)}{\omega(x - n)}\right| = \left|\frac{\sin \omega(x - n)}{\omega(x - n)}\right|^{1-\epsilon} \left|\frac{\sin \omega(x - n)}{\omega(x - n)}\right|^{\epsilon} \leq \frac{1}{\omega^{\epsilon} \, | x - n |^{\epsilon}}.$$

Thus

$$| f(x) | \leq 2M\left\{1 + \frac{1}{\pi\omega^{\epsilon}} \sum_{1}^{\infty} \frac{1}{n^{1+\epsilon}}\right\} \leq 2M\left\{1 + \frac{1}{\pi\omega^{\epsilon}} + \frac{1}{\epsilon\omega^{\epsilon}}\right\}.$$

Now $\epsilon\omega^{\epsilon}$ is a maximum when $1/\epsilon = \log (1/\omega)$ since $0 < \omega < 1$; with this value of ϵ,

$$| f(x) | \leq 2M\{2 + e \log (1/\omega)\}.$$

Substituting $\omega = (\pi - k)/\pi$, we obtain inequality (3).

2. Proof that

$$A(k) \geq \frac{2}{\pi} \log \left\{\frac{4}{\pi}\left[\frac{\pi}{\pi - k}\right]\right\}. \tag{8}$$

The function

$$\begin{aligned} &\frac{1}{2n} \sin n\theta \cot \tfrac{1}{2}\theta \\ &= \frac{1}{n}\{\tfrac{1}{2} + \cos \theta + \cos 2\theta + \cdots + \cos (n - 1)\theta + \tfrac{1}{2} \cos n\theta\} \end{aligned} \tag{9}$$

is a trigonometric polynomial of degree n; at the points $\theta = \nu\pi/n$ it is equal to 1 if $\nu = 0$ or $\pm 2n$, and to 0 if ν is any other integer between $-2n$ and $2n$.

If $\alpha_1, \alpha_2, \cdots, \alpha_{2n}$ are given numbers with $-1 \leq \alpha_\nu \leq 1$, let

$$p(\theta) = \frac{1}{2n} \sum_{\nu=1}^{2n} \alpha_\nu \sin (n\theta - \nu\pi) \cot\left(\frac{\theta}{2} - \frac{\nu\pi}{2n}\right). \tag{10}$$

(This is the interpolation formula of M. Riesz [10].) Then $p(\theta)$ is a trigonometric polynomial of degree n or less, and

$$p(\nu\pi/n) = \alpha_\nu \qquad (\nu = 1, 2, \cdots, 2n).$$

To find the term of n-th degree of $p(\theta)$ we use relation (9), and find that this term is

$$\frac{1}{2n}\sum_{\nu=1}^{2n} \alpha_\nu \cos (n\theta - \nu\pi) = \frac{1}{2n}\sum_{\nu=1}^{2n} (-1)^\nu \alpha_\nu \cos n\theta. \tag{11}$$

Now let

$$\alpha_\nu = \begin{cases} (-1)^{\nu+1} & (\nu = 1, 2, \cdots, n), \\ (-1)^\nu & (\nu = n+1, n+2, \cdots, 2n). \end{cases}$$

Then (11) shows that the term of n-th degree of $p(\theta)$ vanishes, and so $p(\theta)$ is of degree $n - 1$ or less. On the other hand, if these values of α_ν are substituted in (10) we have

$$p\left(\frac{\pi}{2n}\right) = \frac{1}{2n}\sum_{\nu=1}^{n} \cot\left(\frac{\nu\pi}{2n} - \frac{\pi}{4n}\right) - \frac{1}{2n}\sum_{\nu=n+1}^{2n} \cot\left(\frac{\nu\pi}{2n} - \frac{\pi}{4n}\right),$$

and finally

$$p\left(\frac{\pi}{2n}\right) = \frac{1}{n}\sum_{\nu=1}^{n} \cot\left(\frac{\nu\pi}{2n} - \frac{\pi}{4n}\right) > \frac{2}{\pi}\int_{\pi/(4n)}^{\frac{1}{2}\pi} \cot\theta \, d\theta > \frac{2}{\pi}\log\frac{4n}{\pi}.$$

If k is given, $0 < k < \pi$, let n be defined by $n = [\pi/(\pi - k)]$ and let $p(\theta)$ be the trigonometric polynomial of degree $n - 1$ defined above. Then the function $f(z) = p(\pi z/n)$ is an entire function of exponential type $\pi(n-1)/n < k$, and it is bounded by 1 at the integers. But

$$f(\tfrac{1}{2}) = p\left(\frac{\pi}{2n}\right) > \frac{2}{\pi}\log\frac{4n}{\pi} = \frac{2}{\pi}\log\left\{\frac{4}{\pi}\left[\frac{\pi}{\pi - k}\right]\right\}.$$

This relation proves (8).

Bibliography

1. R. P. Boas, Jr., *Entire functions bounded on a line*, this Journal, vol. 6(1940), pp. 148–169.
2. M. L. Cartwright, *On certain integral functions of order one*, Quarterly Journal of Mathematics (Oxford series), vol. 7(1936), pp. 46–55.
3. K. Grandjot, *Ueber Polynome, die in Einheitswurzeln beschränkt sind*, Jahresbericht der Deutschen Mathematiker-Vereinigung, vol. 34(1926), pp. 80–86; vol. 35(1927), p. 112.

4. T. H. Gronwall, *A sequence of polynomials connected with the n-th roots of unity*, Bulletin of the American Mathematical Society, vol. 27(1920-21), pp. 275–279.

5. E. Landau, *Bemerkungen zu einer Arbeit von Herrn Carleman: Ueber die Fourierkoeffizienten einer stetigen Funktion*, Mathematische Zeitschrift, vol. 5(1919), pp. 147–153.

6. A. J. Macintyre, *Laplace's transformation and integral functions*, Proceedings of the London Mathematical Society (2), vol. 45(1938-39), pp. 1–20.

7. A. Pfluger, *On analytic functions bounded at the lattice points*, Proceedings of the London Mathematical Society (2), vol. 42(1936-37), pp. 305–315.

8. G. Pólya, *Prolongement analytique*, L'Enseignement Mathématique, vol. 22(1922), pp. 298–299.

9. G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, Berlin, 1925, vol. 1, p. 116, problem 165.

10. M. Riesz, *Eine trigonometrische Interpolationsformel und einige Ungleichungen für Polynome*, Jahresbericht der Deutschen Mathematiker-Vereinigung, vol. 23(1914), pp. 354–368.

11. G. Valiron, *Sur la formule d'interpolation de Lagrange*, Bulletin des Sciences Mathématiques (2), vol. 49(1925), pp. 181–192, 203–224.

Duke University and Stanford University.

SOME GEOMETRIC INEQUALITIES

By Willy Feller

1. The problem treated in this note was originally formulated by R. Salem and D. C. Spencer in connection with a number-theoretical investigation. Consider a plane domain Γ contained in the unit circle; suppose that the intersection of Γ with any straight line has a measure not exceeding a fixed constant $\delta < 1$. What can be said about the measure M of Γ? If Γ is convex, its area is obviously not greater than $\delta^2\pi$. It is less trivial that, for a convex domain Γ,

$$M \leq \tfrac{1}{4}\delta^2\pi \tag{1}$$

and that the sign of equality holds only if Γ is the interior of a circle; this result was proved by different methods by Bieberbach [1] and Kubota [3]; see also [2], particularly §§44, 54.

It has been widely conjectured that in general $M = O(\delta^2)$ as $\delta \to 0$, or at least that $M = o(\delta)$. Now a simple application of Fubini's theorem shows immediately that necessarily

$$M < 2\delta. \tag{2}$$

It will be proved in the sequel that (2) is the *best* result. In fact, we shall construct a domain Γ (consisting of a finite number of annuli) such that its intersection with any straight line of the plane has a total length not exceeding δ, whereas for its area M we have

$$M > 2\delta(1 - \delta^2\pi^{-2} - \epsilon), \tag{3}$$

where $\epsilon > 0$ is arbitrarily small. Thus 2δ is the best asymptotic estimate for the maximum of the area. The $1/\pi^2$ which multiplies δ^2 is, of course, not the best possible. Our construction can easily be refined, but this seems to be of no interest.

In §4 the above mentioned theorem of Bieberbach and Kubota will be proved in a new simple way which will make the result appear almost trivial. Actually the new proof is even slightly more general.

The generalization of the last result to n dimensions is straightforward. In order to solve the problem of the best estimate in the general case and in n dimensions, we shall (§5) formulate, and solve, a more general and purely analytic problem; it will be seen that our problem actually reduces to an inequality between two integrals.

2. Let $0 < \delta < 1$ be given and denote by N an arbitrarily large but fixed integer. Put

$$N' = [N(1 - \delta^2\pi^{-2})^{\frac{1}{2}}]. \tag{4}$$

Received June 22, 1942.

For $n = 0, 1, \cdots, N' - 1$, we define numbers ρ_n and $r_n > 0$ by

$$\rho_n = \frac{n}{N'},$$

$$r_n^2 - \rho_n^2 = \frac{2\delta}{\pi}\{(1 - \rho_n^2)^{\frac{1}{2}} - (1 - \rho_{n+1}^2)^{\frac{1}{2}}\}. \tag{5}$$

Obviously

$$r_n^2 - \rho_n^2 = \frac{2\delta}{\pi}\frac{\rho_{n+1}^2 - \rho_n^2}{(1 - \rho_n^2)^{\frac{1}{2}} + (1 - \rho_{n+1}^2)^{\frac{1}{2}}} < \frac{\delta}{\pi}\frac{\rho_{n+1}^2 - \rho_n^2}{(1 - \rho_{n+1}^2)^{\frac{1}{2}}}, \tag{6}$$

and hence by (4), for $n \leq N' - 1$,

$$r_n^2 - \rho_n^2 < \rho_{n+1}^2 - \rho_n^2 .$$

It is, therefore, seen that

$$\rho_n < r_n < \rho_{n+1} . \tag{7}$$

Now let r denote the distance from a fixed point O and let A_n stand for the annulus

$$A_n : \qquad \rho_n < r < r_n \qquad\qquad (n = 0, \cdots, N' - 1).$$

By (7) these annuli are non-overlapping. We shall show that, for N sufficiently large, the domain

$$\Gamma = \sum_{n=0}^{N'-1} A_n$$

has all the required properties.

First, it is clear that Γ is contained in the unit circle. Since the A_j are not overlapping, we have for the area M of Γ

$$\begin{aligned} M &= \pi \sum_{n=0}^{N'-1} (r_n^2 - \rho_n^2) \\ &= 2\delta \sum_{n=0}^{N'-1} \{(1 - \rho_n^2)^{\frac{1}{2}} - (1 - \rho_{n+1}^2)^{\frac{1}{2}}\} \\ &= 2\delta\{1 - (1 - \rho_{N'}^2)^{\frac{1}{2}}\}. \end{aligned} \tag{8}$$

Now

$$\begin{aligned} \rho_{N'}^2 &\geq \frac{1}{N^2}(N(1 - \delta^2\pi^{-2})^{\frac{1}{2}} - 1)^2 \\ &\geq 1 - \delta^2\pi^{-2} - \frac{2}{N}, \end{aligned}$$

and hence the inequality (3) follows from (8), provided only that N is sufficiently large.

It remains to prove that the total length of the intersection of Γ with any straight line does not exceed δ.

3. Consider the intersection of a fixed annulus A_k with a straight line whose distance from O is p. If $p \leq \rho_k$, the length of the intersection is

$$\lambda_k = 2\{(r_k^2 - p^2)^{\frac{1}{2}} - (\rho_k^2 - p^2)^{\frac{1}{2}}\},$$

and, in this interval, λ_k increases with increasing p. If $\rho_k < p \leq r_k$, the length of the intersection will be

$$\lambda_k = 2(r_k^2 - p^2)^{\frac{1}{2}},$$

whereas for $p > r_k$ no intersection exists. Hence the maximum of λ_k is attained for $p = \rho_k$ and

$$\max \lambda_k = 2(r_k^2 - \rho_k^2)^{\frac{1}{2}}.$$

Consider now an arbitrary but fixed straight line and let $p < 1$ be its distance from the origin. Let ν be determined by the condition that

$$\rho_{\nu-1} \leq p < \rho_\nu .$$

It is clear that for the total length L of the intersection of this line with Γ we have

$$L \leq 2(r_{\nu-1}^2 - \rho_{\nu-1}^2)^{\frac{1}{2}} + 2(r_\nu^2 - \rho_\nu^2)^{\frac{1}{2}} + 2\sum_{n=\nu+1}^{N'-1} \{(r_n^2 - \rho_\nu^2)^{\frac{1}{2}} - (\rho_n^2 - \rho_\nu^2)^{\frac{1}{2}}\}.$$

Hence

$$\tfrac{1}{2}L \leq (\rho_\nu^2 - \rho_{\nu-1}^2)^{\frac{1}{2}} + (\rho_{\nu+1}^2 - \rho_\nu^2)^{\frac{1}{2}} + \sum_{n=\nu+1}^{N'-1} \frac{r_n^2 - \rho_n^2}{(r_n^2 - \rho_\nu^2)^{\frac{1}{2}} + (\rho_n^2 - \rho_\nu^2)^{\frac{1}{2}}},$$

or, by (6) and the definition of ρ_ν,

$$\tfrac{1}{2}L \leq \left(\frac{2\nu - 1}{N^2}\right)^{\frac{1}{2}} + \left(\frac{2\nu + 1}{N^2}\right)^{\frac{1}{2}} + \frac{\delta}{2\pi} \sum_{n=\nu+1}^{N'-1} \frac{\rho_{n+1}^2 - \rho_n^2}{(1 - \rho_{n+1}^2)^{\frac{1}{2}}(\rho_n^2 - \rho_\nu^2)^{\frac{1}{2}}}$$

(9)

$$= O\left(\frac{1}{N^{\frac{1}{2}}}\right) + \frac{\delta}{2\pi} \sum_{n=\nu+1}^{N'-1} \frac{2n + 1}{(N^2 - (n + 1)^2)^{\frac{1}{2}}(n^2 - \nu^2)^{\frac{1}{2}}}.$$

We shall evaluate the last sum by means of an integral. For a given fixed ν define

$$x_{n-\nu} = \left(\frac{n^2 - \nu^2}{N^2 - \nu^2}\right)^{\frac{1}{2}} \qquad (n = \nu, \nu + 1, \cdots, N' - 1).$$

Obviously, for $\nu < n \leq N' - 1$,

$$0 \leq x_{n-\nu}^2 \leq \frac{N'^2 - \nu^2}{N^2 - \nu^2} \leq 1 - \frac{\delta^2}{\pi^2} \cdot \frac{N^2}{N^2 - \nu^2} < 1 - \frac{\delta^2}{\pi^2}.$$

Next we show that the differences $\Delta x_{n-\nu} = x_{n+1-\nu} - x_{n-\nu}$ tend uniformly to zero as $N \to \infty$. To see this we observe that from

$$2n + 1 = (x_{n+1-\nu}^2 - x_{n-\nu}^2)(N^2 - \nu^2)$$

it follows that

$$0 < x_{n+1-\nu} - x_{n-\nu} = \frac{2n+1}{N^2 - \nu^2} \cdot \frac{1}{x_{n+1-\nu} + x_{n-\nu}}$$

$$\leq \tfrac{1}{2} \frac{2n+1}{(N^2 - \nu^2)x_{n-\nu}} = \tfrac{1}{2} \frac{2n+1}{((N^2 - \nu^2)(n^2 - \nu^2))^{\frac{1}{2}}}.$$

Now $(2n + 1)^2/(n^2 - \nu^2)$ decreases as n increases; the maximum of the last expression is therefore obtained for $n = \nu + 1$, and, accordingly, we have

$$0 < x_{n+1-\nu} - x_{n-\nu} \leq \tfrac{1}{2} \frac{2\nu + 3}{((N^2 - \nu^2)(2\nu + 1))^{\frac{1}{2}}} < 2\left(\frac{\nu}{N^2 - \nu^2}\right)^{\frac{1}{2}}$$

$$< 2\left(\frac{N'}{N^2 - N'^2}\right)^{\frac{1}{2}} < 2\left(\frac{\pi^2}{\pi^2 - \delta^2} \cdot \frac{1}{N'}\right)^{\frac{1}{2}},$$

which is uniformly small as $N \to \infty$. Now

$$\sum_{n=\nu+1}^{N-1} \frac{2n+1}{(N^2 - (n+1)^2)^{\frac{1}{2}}(n^2 - \nu^2)^{\frac{1}{2}}} = \sum_{n=\nu+1}^{N'-1} \frac{x_{n+1-\nu}^2 - x_{n-\nu}^2}{(1 - x_{n+1-\nu}^2)^{\frac{1}{2}} x_{n-\nu}} \tag{10}$$

$$= \sum_{n=1}^{N'-\nu-1} \frac{(x_{n+1} + x_n)\Delta x_n}{x_n(1 - x_{n+1}^2)^{\frac{1}{2}}}.$$

The differences Δx_n are positive and uniformly small; furthermore, $x_0 = 0$ and $x_{N'-\nu-1}$ tends to $1 - \delta^2/\pi^2$ as $N \to \infty$. Hence the sum in (10) tends to the integral

$$2\int_0^{(1-\delta^2\pi^{-2})^{1/2}} \frac{dx}{(1 - x^2)^{\frac{1}{2}}}.$$

It follows that the right member of (9) tends to

$$\frac{\delta}{\pi} \int_0^{(1-\delta^2\pi^{-2})^{1/2}} \frac{dx}{(1 - x^2)^{\frac{1}{2}}},$$

and this is smaller than $\frac{1}{2}\delta$. Thus, by (9), for N sufficiently large, we have

$$L < \delta,$$

which finishes the proof.

4. We turn now to the theorem of Bieberbach and Kubota. Let Γ be the domain bounded by the plane convex curve γ, and let δ be its diameter, that is to say, the maximum length of the intersection of Γ with a straight line. We introduce polar coördinates r, θ around a point P of γ so that Γ lies in the half-plane $0 < \theta < \pi$. If r_θ stands for the radius vector from P to a variable point P_θ of γ, the area M of Γ is given by

$$M = \tfrac{1}{2} \int_0^\pi r_\theta^2 \, d\theta = \tfrac{1}{2} \int_0^{\frac{1}{2}\pi} (r_\theta^2 + r_{\theta+\frac{1}{2}\pi}^2) \, d\theta. \tag{11}$$

Now $r_\theta^2 + r_{\theta+\frac{1}{2}\pi}^2$ equals the square of the distance from P_θ to $P_{\theta+\frac{1}{2}\pi}$. Hence, by assumption,

$$r_\theta^2 + r_{\theta+\frac{1}{2}\pi}^2 \leq \delta^2. \tag{12}$$

The inequality (1) is an immediate consequence of (11) and (12). Moreover, it is seen that the sign of equality in (1) will hold if, and only if,

$$r_\theta^2 + r_{\theta+\frac{1}{2}\pi}^2 = \delta^2, \tag{13}$$

identically. Since P is an arbitrary point of γ, this means that the hypotenuse of any right triangle inscribed in γ has the length δ. It follows immediately that γ is a circle. In fact, let P' be the point on γ for which the chord $P'P_\theta$ is perpendicular to PP_θ. Then $PP_\theta P'$ is a right triangle inscribed in γ, and therefore the chord PP' has the length δ. Now, without loss of generality, we may suppose that γ possesses a tangent at P, that is to say, that $r_\theta \neq 0$ for $0 < \theta < \pi$. It follows then from (13) that $\theta = \frac{1}{2}\pi$ is the only direction for which possibly $r_\theta = \delta$. Hence the point P' has coördinates $r = \delta$ and $\theta = \frac{1}{2}\pi$, and, since P_θ is the projection of P' on the straight line θ, we have $r_\theta = \delta \sin \theta$, which is the equation of a circle.

It will be noticed that our proof of (1) did not fully utilize the assumption that the diameter of Γ is δ, nor the assumed convexity of γ. The proof will hold for any domain Γ bounded by a curve γ for which there exists a point P on γ such that (i) γ has a supporting line at P, and (ii) any chord (with possibly a finite number of exceptions) of γ subtended by a right angle with vertex at P has a length not exceeding δ. In this larger class of domains for which (1) holds, the disc is no longer the only extremal domain. Other examples of extremal domains are the domains defined by $r = \delta$ for $0 < \theta < \frac{1}{2}\pi$, or $r = \delta$ for either $0 < \theta < \pi/4$ or $3\pi/4 < \theta < \pi$ and $r = 0$ for other values of θ.

The generalization of this result to the case of n dimensions is straightforward.

5. In order to solve our general maximum problem for n dimensions we shall reformulate it in analytical language.

Let $f(x, y)$ be the characteristic function of the domain Γ, that is, let $f(x, y) = 1$ if $(x, y) \in \Gamma$, and $f(x, y) = 0$ otherwise. Then the measure M of Γ is defined by

$$M = \iint_{-\infty}^{+\infty} f(x, y)\, dx\, dy; \tag{14}$$

the assumption that the measure of the intersection of Γ with any straight line does not exceed δ means that

$$\int_{-\infty}^{+\infty} f(x + t \cos \alpha, y + t \sin \alpha)\, dt \leq \delta \tag{15}$$

for all x, y and α; the integrals in (14) and (15) are, of course, proper integrals since $f(x, y)$ vanishes outside of the unit circle. We shall now generalize our problem by dropping the assumption that $f(x, y)$ is the characteristic function of a set. Thus we require *the least upper bound of the quantity M defined by* (14), *when $f(x, y)$ is any non-negative measurable function vanishing for $x^2 + y^2 \geq 1$ and such that* (15) *holds.*

In order to solve this problem (and its analogue in n dimensions) we shall first show that *it suffices to consider functions $f(x, y)$ depending on $r = (x^2 + y^2)^{\frac{1}{2}}$ only.* In fact, for a given $f(x, y)$, define a new function $F(x, y)$ by

$$F(x, y) = \frac{1}{2\pi} \int_0^{2\pi} f(x \cos \theta + y \sin \theta, - x \sin \theta + y \cos \theta)\, d\theta. \tag{16}$$

Obviously $F(x, y)$ is again non-negative and $F(x, y) = 0$ for $x^2 + y^2 \geq 1$. Furthermore, it is readily seen that, for any x, y and α,

$$\begin{aligned} \int_{-\infty}^{+\infty} & F(x + t \cos \alpha, y + t \sin \alpha)\, dt \\ &= \frac{1}{2\pi} \int_0^{2\pi} d\theta \int_{-\infty}^{+\infty} f(x \cos \theta + y \sin \theta + t \cos (\theta - \alpha), \\ &\qquad\qquad - x \sin \theta + y \cos \theta + t \sin (\theta - \alpha))\, dt \\ &\leq \frac{1}{2\pi} \int_0^{2\pi} \delta\, d\theta = \delta, \end{aligned} \tag{17}$$

whereas for the "measure" of $F(x, y)$ we obtain

$$\iint_{-\infty}^{+\infty} F(x, y)\, dx\, dy = \frac{1}{2\pi} \int_0^{2\pi} d\theta \iint_{-\infty}^{+\infty} f(\xi, \eta)\, d\xi\, d\eta = M. \tag{18}$$

Accordingly, $F(x, y)$ satisfies all the conditions imposed on $f(x, y)$ and has the same "measure"; we can therefore consider $F(x, y)$ instead of $f(x, y)$. However, $F(x, y)$ obviously depends only on r:

$$F(x, y) = \Phi(r), \qquad r = (x^2 + y^2)^{\frac{1}{2}}. \tag{19}$$

The "measure" (18) of $F(x, y)$ can, therefore, be written in the form

$$M = 2\pi \int_0^1 \Phi(r) r \, dr; \tag{20}$$

as for the integral (17), it suffices for reasons of symmetry to consider the case where $x = p$ $(0 \leq p < 1)$, $y = 0$, and $\alpha = \frac{1}{2}\pi$. The integral (17) then reduces to

$$\int_{-\infty}^{+\infty} F(p, t) \, dt = 2 \int_0^{\infty} \Phi((p^2 + t^2)^{\frac{1}{2}}) \, dt,$$

and condition (15) now reads:

$$\int_p^1 \Phi(r) \frac{r}{(r^2 - p^2)^{\frac{1}{2}}} \, dr \leq \tfrac{1}{2}\delta. \tag{21}$$

In the case of n dimensions there is no change in (21); the integral (20) is to be replaced by

$$M = \omega_{n-1} \int_0^1 \Phi(r) r^{n-1} \, dr, \tag{22}$$

where ω_{n-1} stands for the volume of the $(n - 1)$-dimensional unit sphere. Our general problem is thus reduced to that of *finding the least upper bound of the integral* (22) *for all non-negative functions* $\Phi(r)$ *for which* (21) *holds for any* $0 \leq p < 1$.

Now from (21) we obtain

$$\int_0^1 p^{n-2} \, dp \int_p^1 \Phi(r) \frac{r}{(r^2 - p^2)^{\frac{1}{2}}} \, dr \leq \tfrac{1}{2}\delta \int_0^1 p^{n-2} \, dp = \frac{\delta}{2(n - 1)};$$

reversing the order of integration and observing that

$$\int_0^r \frac{p^{n-2}}{(r^2 - p^2)^{\frac{1}{2}}} \, dp = r^{n-2} \int_0^{\frac{1}{2}\pi} \cos^{n-2}\lambda \, d\lambda = r^{n-2} \frac{\omega_{n-1}}{2\omega_{n-2}}, \tag{23}$$

we obtain

$$\frac{\omega_{n-1}}{2\omega_{n-2}} \int_0^1 \Phi(r) r^{n-1} \, dr \leq \frac{\delta}{2(n - 1)},$$

or using (22),

$$M \leq \frac{\delta\omega_{n-2}}{(n - 1)} = \delta V_{n-1}, \tag{24}$$

where V_{n-1} is the volume of the interior of a $(n - 1)$-dimensional sphere in n-dimensional space. (In order that (23) be valid also for $n = 2$ we define

$\omega_0 = 2$.) Here the sign of equality will hold if, and only if, the sign of equality holds in (21) for almost all p. Now let

$$\Phi(r) = \frac{\delta}{\pi} \frac{1}{(1 - r^2)^{\frac{1}{2}}}; \tag{25}$$

then the sign of equality holds in (21) for all p, as is seen upon introducing the substitution

$$r^2 = p^2 + (1 - p^2) \sin^2 \lambda.$$

Hence (24) *is the best estimate for* M, and our generalized problem is solved. The same estimate is also the best in the case of characteristic functions of sets; this is seen by approximating the function $\Phi(r)$ of (25) by a step function, assuming the values 0 and 1 only.

Bibliography

1. L. Bieberbach, *Ueber eine Extremaleigenschaft des Kreises*, Jber. Deutsche Math.-Verein., vol. 24(1915), pp. 247–250.
2. T. Bonnesen and W. Fenchel, *Theorie der konvexen Körper*, Ergebnisse der Math., vol. 3, Berlin, 1934.
3. T. Kubota, *Ueber die konvex-geschlossenen Mannigfaltigkeiten im n-dimensionalen Raume*, Science Reports Tôhoku University, vol. 14(1925), pp. 85–99.

Brown University.

A PARTIAL SOLUTION OF A PROBLEM OF J. R. KLINE

By Dick Wick Hall

In this paper we consider a compact and metric locally connected continuum M satisfying the following conditions: (a) no two points separate M, (b) every simple closed curve J of M separates M, (c) for any simple closed curve J of M the number of components of $M - J$ is finite. E. E. Betz [1] has, without the use of (a), solved certain accessibility problems arising in such spaces. But whether or not M is homeomorphic with a 2-sphere is an unsolved special case of a well-known problem of J. R. Kline. The purpose of this paper is to give an affirmative solution to this problem. It seems highly suggestive that the condition (c) is not required in most of our lemmas. Inasmuch as our theorem might be true without this condition, we have indicated the lemmas in which it is not required by placing a prime after their numbers.

Theorem. *In order that a locally connected continuum M be homeomorphic with a 2-sphere it is necessary and sufficient that it satisfy the following conditions:*

(a) *no two points separate M,*

(b) *every simple closed curve J of M separates M,*

(c) *for any simple closed curve J of M the number of components of $J - M$ is finite.*

Proof. Since the given conditions are obviously necessary we content ourselves with a demonstration of their sufficiency. We note that M has no cut-point. Hence for any point p of M there exists an arbitrarily small region R about p having property S and such that $M - R$ is connected. Whenever we take a region about a point we shall assume that it has both of these properties.

We begin by dividing the points of M into two mutually exclusive classes, so that we can write $M = B + I$. A point b of M is said to be a *regular* point of M provided there exists a region R in M containing b and homeomorphic with an open 2-cell. The set B consists of all regular points of M, and I is defined by the equation above. We shall call the points of I the *irregular* points of M. We have at once:

(i′) *B is an open set, and every component of B is open.*

(ii′) *Every simple closed curve J of M contained in B separates M into exactly two components and is the complete boundary of each of them.*

Received June 23, 1942; presented to the American Mathematical Society, September 8, 1942. This paper resulted from considerations started when the author was a National Research Fellow at the University of Pennsylvania. The author wishes to take this opportunity to express his deep appreciation for many helpful conversations with Professor Kline on the general problem of which this paper presents a partial solution. He is also grateful for several suggestions from F. B. Jones.

To see this let p be any point of J. Then, since p is regular, there exists a region R in M containing p and homeomorphic with an open 2-cell. We know that J separates this open 2-cell into exactly two components. Moreover, every component of $M - J$ having p as a limit point must have points in one of the two components of $R - J$. Consequently, there is an open arc of J having p as an interior point and consisting entirely of boundary points of every such component. It follows at once that this implies the existence of precisely two components E and F of $M - J$ and that every point of J must be a limit point of both E and F. This proves (ii′).

It is easily seen that (ii′) remains true if we replace M throughout by G, any component of B. Hence we have (see [3; Theorem V′]; this paper also contains a summary of the previous work and an excellent bibliography)

(iii′) *Every component G of B is homeomorphic with a region on a sphere. The complement of this region on the sphere can be taken as a totally disconnected point set.*

Using (iii′) we easily deduce:

(iv′) *If T, K, and L are three arcs in B disjoint except for their common end points p and q, then $M - (T + K + L)$ is the sum of three components having $T + K$, $K + L$, and $L + T$ as their boundaries.*

(v′) *I is vacuous or a perfect set.*

We know by (i′) that I is closed. Hence if I is not perfect there must exist an isolated point w of I. Let R be a region in M containing w such that the locally connected continuum $\bar{R}$ is disjoint with $I - w$.

We shall show first that $\bar{R}$ is homeomorphic with a subset of a cactoid (a locally connected continuum every true cyclic element of which is homeomorphic with a sphere). To do this it suffices, by a theorem of Claytor [2], to prove that $\bar{R}$ contains neither of the two primitive skew curves $\mathfrak{C}$ or $\mathfrak{D}$ of Kuratowski. We recall what these two primitive skew curves are. The curve $\mathfrak{C}$ is homeomorphic with a complex consisting of two groups of three vertices each and nine 1-cells, in a fashion that each vertex of one group together with each vertex of the other group bounds a 1-cell. The curve $\mathfrak{D}$ consists of five vertices and ten 1-cells, in a fashion that each pair of vertices bounds a 1-cell. By $\mathfrak{H}$ we denote either of the primitive skew curves or their homeomorphs.

Suppose that $\mathfrak{H}$ is contained in $\bar{R}$. We may evidently assume that $\mathfrak{H}$ is contained in R since if it were not we could take a smaller region and find a new primitive skew curve. Now $R - w$ consists entirely of regular points and hence by (iii′) every component of $R - w$ is homeomorphic with a region on a sphere. It is immediate that w is a point of $\mathfrak{H}$. Let T be a theta curve in $\mathfrak{H}$ consisting of three independent arcs $aa'b$, $ab'b$, $ac'b$ and not containing the point w.

We know by (iv′) that T separates M into precisely three complementary domains A', B', C' having the three simple closed curves $ab'bc'a$, $aa'bc'a$, and $aa'bb'a$, respectively, as their complete boundaries. Now suppose $\mathfrak{H}$ is $\mathfrak{C}$, and

that $\mathfrak{C}$ consists of the two groups of vertices (a, b, c), (a', b', c'). We have at once that c must lie in either A', B', or C'. Consequently, c cannot be joined to all three of the points a', b', c' by arcs in M which fail to intersect T. This contradiction shows us that $\mathfrak{H}$ cannot be $\mathfrak{C}$.

Finally, assume that $\mathfrak{H}$ is $\mathfrak{D}$. We may assume that the five vertices defining $\mathfrak{D}$ are a', b', c', a, and b. Now $\mathfrak{D}$ is symmetric in its five points, hence we need to consider only two cases. Assume first that w is an interior point of some arc of $\mathfrak{D}$. We lose no generality if we assume that w lies on the arc of $\mathfrak{D}$ joining a and b and containing none of the points a', b', c'. It is easily seen that the simple closed curve $a'b'c'a'$ of $\mathfrak{D}$ separates M between a and b. Consequently, it is impossible to find the arc awb of $\mathfrak{D}$. We may thus assume that w is a vertex of $\mathfrak{D}$, say $w = b'$. Let z be an interior point of the arc ab of $\mathfrak{D}$ and T' the theta curve in $\mathfrak{D}$ consisting of the three arcs azb, $aa'b$, $ac'b$. We know that $M - T'$ has exactly three components A', C', Z having the three simple closed curves $azbc'a$, $azba'a$, $aa'bc'a$, respectively, as their complete boundaries. Now b' can be joined to both c' and a' by arcs of $\mathfrak{D}$. Thus b' lies in Z. For the same reason the arc $a'c'$ of $\mathfrak{D}$ lies in Z. It is easily seen that $a'c'$ separates $\bar{Z}$ between b' and either a or b so that b' cannot be joined to both these points in $M - aa'c'a$. This contradiction shows us that $\bar{R}$ cannot contain either $\mathfrak{C}$ or $\mathfrak{D}$ and establishes our assertion that this locally connected continuum is homeomorphic with a subset of a cactoid.

We may thus assume that $\bar{R}$ is imbedded in a cactoid Y. Now no two points of M separate M. Hence no nondegenerate acyclic [4] element of Y can intersect a sufficiently small neighborhood of w. Therefore, w must lie on a 2-sphere S which is a true cyclic element of Y, and we let L be the part of $\bar{R}$ lying on S. Now w must be a limit point of $S - L$. Otherwise we can find a 2-cell Z in L having w as an interior point. But w lies in I and hence w must be a limit point of $M - L$. It follows easily that for a sufficiently small simple closed curve J in Z surrounding w the set $M - J$ is connected. This contradiction shows us that w must be a limit point of $S - L$. Thus w lies on a simple arc A of L which is composed of boundary points of some complementary domain of L in S, and every point of $A - w$ is a point of B. It is then a simple matter to find a simple closed curve in B which is not the complete boundary of both of its complementary domains. This contradiction completes the proof of (v′).

(vi′) *If p is any point of M and R is any region in M containing p, then p is not a limit point of cut-points of the locally connected continuum $K = \bar{R}$.*

We know that p is not an end-point of K since M is cyclic, and we have seen that no nondegenerate acyclic element of K can intersect a sufficiently small neighborhood of p. Consequently, p lies in a true cyclic element E of K. Now every component of $K - E$ contains a limit point of $M - K$ since M is cyclic. Thus, since the components of $K - E$ are either finite in number or form a null sequence, it follows that p cannot be a limit point of cut-points of K which

lie in E. Let $\{w_i\}$ be a sequence of cut-points of K converging to p. By the reasoning above we see that no generality is lost if we assume that all the w_i lie in a single component Z of $K - E$. It follows at once that p is the boundary of Z in the locally connected continuum K, hence p is a cut-point of K. Now every w_i is a cut-point of the locally connected continuum $\bar{Z}$. Hence, since p is a non-cut-point of $\bar{Z}$ it follows that p is an end-point of $\bar{Z}$, which result gives an immediate contradiction to (a).

(vii) *If p is any point of M and R is any region in M containing p, then there exists a region R' in M containing p such that for any continuum F in $R' - p$ no two points separate the locally connected continuum $K = \bar{R}$ between p and F.*

(vii′) *In the proof of* (vii) *hypothesis* (c) *of the theorem may be replaced by the condition that no three points separate M.*

We shall prove (vii) and (vii′) simultaneously. Assume the conclusion false, and let R_1 be a region in M containing p, of diameter less than one, and sufficiently small so that it contains no cut-point of K distinct from p. Then there exists a continuum F_1 in $R_1 - p$ and a set H_1, consisting of one or two points, which separates K between p and F_1. But F_1 lies in R_1 and hence in the same true cyclic element of K as p. Consequently, H_1 consists of two distinct points a_1 and b_1.

Denote by R_1' a region in M containing p sufficiently small so that the locally connected continuum $K_1 = \bar{R}_1'$ is contained in R_1 and contains neither H_1 nor any point of F_1. Pick a region R_2 in M containing p, of diameter less than $\frac{1}{2}$, and sufficiently small so that R_2 contains no cut-point of K_1 distinct from p. Then there exists a continuum F_2 in $R_2 - p$ and a set H_2 which separates K between p and F_2. Since F_2 lies in the same true cyclic element of K_1 as p, it follows that H_2 lies in this same true cyclic element. Consequently, H_2 lies in $\bar{R}_1'$ and consists of two distinct points a_2 and b_2.

Assume that the sets R_i, H_i, and F_i have been defined. Denote by R_i' a region in M containing p and sufficiently small so that the locally connected continuum $K_i = \bar{R}_i'$ is contained in R_i and contains no point of either H_i or F_i. Pick a region R_{i+1} in M containing p, of diameter less than $1/(i + 1)$, and sufficiently small so that it contains no cut-point of K_i distinct from p and has a closure contained in R_i'. Then there exists a continuum F_{i+1} in $R_{i+1} - p$ and a set H_{i+1} which separates K between p and F_{i+1}. But F_{i+1} lies in the same true cyclic element of K_i as p. Hence H_{i+1} lies in this same true cyclic element. It follows at once that H_{i+1} consists of two distinct points a_{i+1} and b_{i+1} each of which lies in K_i. Thus we have at once:

(A) *For every i the set H_i consists of two distinct points a_i and b_i. The a_i are all distinct, the b_i are all distinct, no a_i coincides with any b_i, and each of these sequences converges to p.*

We know that $K - p$ has at most a finite number of components and hence

lose no generality if we assume that there exists a component Z of $K - p$ which contains every H_i and every F_i. Then, since p is not a cut-point of the locally connected continuum $\overline{Z}$ we can assume:

(B) *For every i the set H_i lies in the component of $K - H_j$ containing H_1 for any j greater than i and in the component of this set containing p for any j less than i. All the sets H_i lie in the same true cyclic element of K as p.*

We can now assert:

(C) *By a proper choice of the sets H_i the set $L_i = M - (p + H_i)$ will be disconnected for all i.*

Otherwise, we may assume that L_i is connected for every i. Define $Q = p + \sum H_i$. Then Q is a closed set. Hence $K - Q$ has infinitely many components which form a null sequence. It follows that if V is a neighborhood of p whose closure is contained in R we may assume that H_i (for every i) together with every component of $K - Q$ except one is contained in V. Let D be the component of $M - H_i$ which contains p. It follows that if j be any integer greater than i, then $D - H_j$ is separated, and has one component contained in V. Call this component D'. It is easily seen that if q be a point of D', then any arc joining q to a point of $M - K$ must contain either a point of H_i or a point of H_j for every j greater than i. Consequently, every such arc must contain either p or a point of H_i. This establishes (C) and gives (vii′) as an immediate corollary.

Now let E be the true cyclic element of K containing all the H_i and let p' be a point of E in a component of $E - H_1$ not containing p. Then there exists a simple closed curve J of E containing both p and p'. It is immediate that J contains Q so that $M - J$ has infinitely many components, contradicting (c) of the theorem. This contradiction completes the proof of (vii).

(viii) *If $\{x_i\}$ is an infinite sequence of points of M which converges to a non-local-separating point of M, then there exists a simple closed curve J in M containing infinitely many of the points x_i.*

(viii′) *We get the same conclusion if we replace hypothesis* (c) *by the condition that no three points separate M.*

We again give the proofs simultaneously. To this end let $\{x_i\}$ be any infinite sequence of points of M converging to a point p of M. We must show that there exists a simple closed curve J in M containing infinitely many points of this sequence. We begin by picking a region R_1 in M containing p and of diameter less than one. No generality is lost if we assume that R_1 contains x_i for every i. Define R_1' as the region in R_1 given by (vii) or (vii′). Apply these lemmas again to find the region R_2 in R_1'. We may evidently assume that the diameter of R_2 is less than $\frac{1}{2}$. No generality is lost if we assume that x_1 lies in $M - R_1'$ while R_2 contains every x_i for i greater than one. Now suppose R_{i-1} has been defined. Then R_{i-1} contains every x_j for j greater than $i - 2$, while the remaining x_j lie

in $M - R'_{i-2}$. Choose R'_{i-1} by (vii) or (vii′) and apply one of these lemmas again to find R_i contained in R'_{i-1}. We may assume without loss of generality that the diameter of R_i is less than $1/i$, that R_i contains every x_j for j greater than $i - 1$, and that x_{i-1} lies in $M - R'_{i-1}$.

Now no two points separate M. Hence there exists a simple closed curve J_2 in $M - p$ containing x_1 and x_2. Suppose that we have found a simple closed curve J_{i-1} in $M - p$ containing all the points x_j for j less than i. We lose no generality if we assume that J_{i-1} contains no x_j for j greater than $i - 1$. Now x_{i-1} lies in R_{i-1}. Let $a_{i-1}x_{i-1}b_{i-1} = T_{i-1}$ be an arc of J_{i-1} lying entirely in R_{i-1}. By construction of this region no two points separate $\overline{R}'_{i-2}$ between p and T_{i-1}. Furthermore no two points separate this locally connected continuum between p and x_i. Consequently, no two points separate $\overline{R}'_{i-2}$ between x_i and T_{i-1}. It follows that there exist three independent arcs $x_ih_jx_{i-1}$ ($j = 1, 2, 3$) in $\overline{R}'_{i-2}$ and we let y_j be the first intersection of the arc $x_ih_jx_{i-1}$ with J_{i-1}. There also exist three independent arcs $x_ik_ja_{i-1}$ ($j = 1, 2, 3$) and we let z_j be the first intersection of the arc $x_ik_ja_{i-1}$ with J_{i-1}.

Suppose first that all three of the y_j coincide, i.e., $y_1 = y_2 = y_3 = x_{i-1}$. We may suppose that p lies on none of the arcs $x_ih_2x_{i-1}$, $x_ih_3x_{i-1}$, x_iz_2, x_iz_3, and that x_iz_3 does not contain the point x_{i-1}. If the arc x_iz_3 is disjoint with either of the first two arcs of the preceding sentence, then we easily obtain a simple closed curve J_i having the required properties. Otherwise x_iz_3 must intersect both of these arcs and we let w_3 be its last intersection with their sum. The simple closed curve J_i in $M - p$ is then easily obtained.

Next we suppose that $y_1 = y_2 = x_{i-1}$, while y_3 is distinct from x_{i-1}. If p does not lie on the arc x_iy_3 we get the simple closed curve J_i at once. If p does lie on this arc we can bypass it in an arbitrarily small region and again get J_i. Finally, J_i may be easily obtained if the y_i are all distinct by making use of the fact that p is a non-local-separating point of M.

It is immediate that if we define J as the product of all the J_i, then J is a simple closed curve in M containing infinitely many of the points x_i. Thus we have established both (viii) and (viii′).

(ix′) *The set of all local separating points of M is countable.*

This is immediate from (a) of the theorem and the well-known fact [5] that all save at most a countable number of the local separating points of M are of Menger order 2 in M.

(x) *M is homeomorphic with a sphere.*

By (iii′) it suffices to show that $I = 0$. Assume that this is not the case. Then by (v′) I is a perfect set, and hence all save at most a countable number of points of I are condensation points of this set. It is thus easy, by (ix′) to pick an infinite sequence $\{p_i\}$ of non-local-separating points of M all of which lie in I and which converge to a non-local-separating point p of M. By (viii) we lose

no generality if we assume that all the points p_i lie on a simple closed curve J in M and converge monotonically to p on J.

For each $i = 1, 2, 3, \cdots$, there exists a region R_i in M satisfying the following conditions: (1) $R_i \cdot J$ is an open arc of J, (2) $R_1, R_2, \cdots$ converges to p, (3) $R_i \cdot R_j = 0$ for $i \neq j$. Let i denote a fixed positive integer and K_i the locally connected continuum $\overline{R}_i$. By (vii) there exists in R_i a region R_i' containing p_i such that for any continuum F in $R_i' - p_i$ no two points separate the locally connected continuum K_i between p_i and F. We lose no generality if we assume that R_i' satisfies (1). Now I is a perfect set and hence p_i must be a limit point of points of I lying in R_i'. Suppose there exists no point of I in $R_i' - J$. Then we can find a sequence of points $\{b_i\}$ on J, all belonging to I, all lying in R_i', and such that this sequence converges monotonically to p_i on J. We know that $M - J$ has but a finite number of components, and hence any arc of J must contain infinitely many points which are boundary points of some complementary domain D of $M - J$. This means that for a sufficiently large i we can replace an arc of J containing a point b_i by an arc in D and thus the modified simple closed curve J is such that the point b_i lies in R_i' but not on J, while p_i still lies on J. Consequently, we lose no generality if we assume that there exist points of I in R_i' which are not on J.

Let b be such a point of I and R_i'' a region in M containing b and sufficiently small so that its closure is disjoint with J. Then there exists a simple closed curve J_i in R_i'' which is not the complete boundary of every one of its complementary domains. Now J lies in the component N of $M - J_i$ which contains $M - R_i''$. Let H be any component of $M - J_i$ distinct from N. Then H lies in R_i'' and hence in R_i'. We define $x_iy_iz_i$ as a minimal arc of J_i separating the space M, and let $x_iw_iz_i$ be the complementary arc of J_i.

We shall show that there exists an arc m_in_i lying in K_i and spanning J so that the closure of one of its complementary domains is a subset of K_i. To this end let R_i^3 be a region in M containing p_i but sufficiently small so that its closure is disjoint with both $\overline{R}_i''$ and $M - R_i'$. Now p_i is a non-local-separating point of M and hence cannot be an isolated point of the boundary of any component of $R_i^3 - J$. Let D be the component of this set contained in the component of $K_i - J$ which contains R_i''. Then there exists an arc $a_iv_ic_i$ having its interior in D and spanning J. Let $a_id_ic_i$ be the arc of J in R_i^3 and denote the sum of the two arcs $a_id_ic_i$ and $a_iv_ic_i$ by J_i'. Then there exists a component H' of $M - J_i'$ lying in R_i^3. Pick a point m in H' and a point w in H. Suppose that there exist two points h and k whose sum separates K_i between m and w. Neither of these points can be p_i since this point is a non-local-separating point of M. Consequently, the point pair (h, k) must separate K_i between p_i and one of the two points m, w. This is impossible. It follows that there must exist three independent arcs in K_i joining m to w. We let e_i, f_i, g_i be the last points of intersection of these arcs with $J_i' + J_i$ and j_i, k_i, t_i the first points of intersection of these arcs with J_i following e_i, f_i, g_i. By construction the six points $e_i, f_i, g_i, j_i, k_i, t_i$ are all distinct. The proof now splits into two cases.

Case 1. *There exists a component of $M - J_i$ which does not contain J and which does not have J_i as its complete boundary.*

In this case we lose no generality if we assume that J_i is not the complete boundary of H, and that the arc $x_iy_iz_i$ of J_i is a minimal arc of this simple closed curve containing the boundary of H. By a theorem of Betz [1] $H + x_i + z_i$ contains an arc $x_iq_iz_i$. If two of the points j_i, k_i, t_i lie in the closed arc $x_iw_iz_i$, then the existence of the arc m_in_i is immediate. Hence we may assume that j_i and k_i are both interior points of $x_iy_iz_i$. If either of the arcs $x_iw_iz_i$, $x_iq_iz_i$ separate M the existence of m_in_i again follows at once. Consequently, we may assume that neither of these arcs separates the space. This gives the easy contradiction that the simple closed curve consisting of the sum of these two arcs fails to separate M, and thus proves that an arc m_in_i of the required type must exist.

Case 2. *No such component exists.*

Let Z_i be the component of $M - J_i$ containing J. We lose no generality if we assume that the arc $x_iy_iz_i$ of J_i is the minimal arc of this simple closed curve containing the boundary of Z_i. It follows easily as before that no component of $M - J_i$ has its complete boundary in the closed arc $x_iw_iz_i$. By the theorem of Betz we may find an arc $x_ir_iz_i$ in R_i' having its interior in Z_i. If this spanning arc has two or more points in common with $J_i' + J$ we easily find m_in_i.

Assume next that the arc $x_ir_iz_i$ has precisely one point, which we denote by r_i, in common with $J + J_i'$. We suppose first that r_i lies on J_i' and in this case lose no generality in assuming that r_i is an interior point of $a_iv_ic_i$. Now consider the three arcs e_ij_i, k_if_i, t_ig_i. We may assume that r_i is not on either of the arcs k_if_i or t_ig_i. If either of these arcs has a point in common with $x_ir_iz_i$, the construction of the arc m_in_i is easy. Hence we may assume that the arc $x_ir_iz_i$ is disjoint with both of the arcs k_if_i and g_it_i. If either of the arcs separates the space, we again get m_in_i easily. If, on the other hand, this does not happen, then the sum of these two arcs is a simple closed curve not separating the space. The remaining case is where the point r_i lies on $J - J_i'$. We may assume in this case that the four points r_i, a_i, d_i, c_i lie on the simple closed curve J in this order. We again get that $x_ir_iz_i$ is disjoint with both k_if_i and g_it_i. An arc $a_i'v_i'c_i'$ can now be constructed spanning J in an arbitrarily small region containing r_i and such that r_i lies between a_i' and c_i' on J. If $x_ir_iz_i$ is disjoint with this new spanning arc, we can make a slight modification of J so that this does not occur. If $x_ir_iz_i$ intersects this new spanning arc, we can easily modify $x_ir_iz_i$ so that it meets $J + J_i'$ in more than one point. The remaining case, where $x_ir_iz_i$ is disjoint with $J + J_i'$, follows by precisely the same reasoning as the case where r_i is on J_i. Hence, in every case, we can find the arc m_in_i.

Now if we add to J all of the arcs m_in_i just constructed, the resulting set obviously contains a simple closed curve J' such that $M - J'$ has infinitely many components. This contradiction completes the proof of the theorem.

Bibliography

1. E. E. Betz, *Accessibility and separation by simple closed curves*, American Journal of Mathematics, vol. 63(1941), pp. 127–135.
2. S. Claytor, *Topological immersion of Peanian continua in a spherical surface*, Annals of Mathematics, vol. 35(1934), pp. 809–835.
3. E. R. van Kampen, *On some characterizations of 2-dimensional manifolds*, this Journal, vol. 1(1935), pp. 74–93.
4. A. D. Wallace, *The acyclic elements of a Peano space*, Bulletin of the American Mathematical Society, vol. 47(1941), pp. 778–780.
5. G. T. Whyburn, *Local separating points of continua*, Monatshefte für Mathematik und Physik, vol. 36(1929), pp. 305–314.

Brown University.

CONTENTS